A CT/HoTT Visual Cortex: Seeing Polyhedral Phase Changes and Human Faces as Paths, Charts, and Coherence

Douglas C. Youvan

doug@youvan.com

www.youvan.ai

December 28, 2025

Human vision treats faces, edges, corners, and "parts" as primary. From that stance, some geometric facts feel metaphysically hard: a vertex is attached to three faces or it is attached to four; the cube and octahedron are duals, but the "valence switch" looks like a discontinuity rather than a smooth transition. This paper asks what kind of sentient perception would experience such changes as ordinary. We develop a speculative but disciplined model: a visual cortex whose native primitives resemble category theory and homotopy type theory (HoTT). In this model, perception is functorial rather than snapshot-based; identity is not a Boolean tag but a space of witnessed paths; and what humans call "faces" and "vertices" are late discretizations of a deeper coherence field. We propose that a CT/HoTT-wired mind would perceive combinatorial "phase changes" as chart changes across stratified moduli, not as breaks in being. We then extend the same framework to ordinary objects—especially human faces—arguing that recognition becomes transport along identity paths, while deception (e.g., incoherent motion, mismatched constraints) appears as an immediate failure of lift and coherence. The goal is not to mystify vision, but to compare perceptual metaphysics: object-first versus equivalence-first seeing.

Keywords: category theory, homotopy type theory, perception, visual cortex, duality, cube octahedron, polyhedral phase change, stratified geometry, moduli space, chart change, identity as path, higher morphisms, coherence, incidence, vertex valence, implicit surfaces, face perception, recognition, deepfakes, phenomenology.

A collaboration with GPT-5.2-Thinking. 54 pages. CC4.0.

Outline

1. Problem Statement: When Discrete Structure Feels Metaphysically Hard
1.1 The cube–octahedron duality as a perceptual stress test
1.2 Why “3 faces vs 4 faces” resists smooth visualization in object-first vision
1.3 The paper’s wager: change the perceptual primitives, not the geometry

2. Two Kinds of Seeing: Object-First Versus Relation-First
2.1 Object-first vision: edges, parts, segmentation, stable labels
2.2 Relation-first vision: constraints, adjacency tendencies, coherence fields
2.3 What “ordinary” means in each metaphysics of perception

3. A Minimal Technical Backbone (Kept Intuitive, Not Formalist)
3.1 Categories as “composable perceptions” rather than static images
3.2 Functorial perception: structure-preserving mapping from sensation to shape
3.3 HoTT’s core move: identity as paths, not equality as a switch

4. Polyhedral Phase Change Reframed
4.1 Valence as a derived census, not a primitive fact
4.2 Stratified spaces: smooth regions plus boundary events as normal loci
4.3 Degeneration moments: where a new incidence becomes fully legible
4.4 Duality as binocular perception: primal and dual co-present

5. What a CT/HoTT-Wired Visual Cortex Would Represent
5.1 Percepts as types with families of presentations (charts)
5.2 Attachments as graded participation (soft incidence)
5.3 Higher morphisms as perceptual continuity of re-factorizations
5.4 Coherence as a felt invariant: agreement across multiple parsing routes

6. Phenomenology: What It Would Feel Like to “See” the 3→4 Switch Smoothly
6.1 The “chart change” qualia: re-description without ontological rupture
6.2 The “lift succeeds” qualia: smooth transport of local structure
6.3 The “critical event” qualia: boundary crossing as ordinary perception
6.4 Why humans experience a “snap” and the CT/HoTT mind does not

7. Ordinary World Competence: Seeing Human Faces
7.1 A face as a type: viewpoint, lighting, expression as internal variation
7.2 Recognition as transport: identity witnessed by continuous deformation
7.3 Features as subobjects-by-role: eyes and mouth as relational invariants
7.4 Expression as higher morphism: microexpressions as small homotopies

8. Error, Deception, and the Uncanny
8.1 Coherence breaks as immediate perceptual alarms
8.2 Deepfakes and “failure of lift”: when motion doesn’t match constraints
8.3 Disguises as chart tricks: what changes and what cannot
8.4 Hallucination risk in equivalence-first seeing: false witnesses and overtransport

9. Cognitive Tradeoffs and Limits
9.1 Costs of coherence-first perception: time, compute, and attentional load
9.2 Social metaphysics: individuality as a section through a continuous space
9.3 When discretization is necessary: action, language, and shared reference
9.4 Pathologies: overidentification, undersegmentation, and missed saliencies

10. Testable Predictions and Thought Experiments
10.1 Behavioral signatures: invariance profiles under occlusion and lighting
10.2 Ambiguity handling: multiple identifications retained rather than collapsed
10.3 Detection of synthetic faces: constraint-consistency sensitivity
10.4 The “polyhedron movie” task: graded incidence perception in humans and models

11. Implications for AI, XR, and Neuroscience
11.1 Building perception stacks that are functorial by design
11.2 Representation choices: meshes versus implicit fields versus path-based models
11.3 XR as training for chart-change perception: can humans learn this qualia?
11.4 Ethics: enhanced recognition, privacy, and the temptation of “oracle sight”

12. Conclusion: A Different Ordinary
12.1 Restating the thesis: phase changes are chart changes under coherence-first vision
12.2 Cube–octahedron duality as a parable for everyday recognition
12.3 Forward work: formalization boundaries, experiments, and responsible deployment

References

1. Problem Statement: When Discrete Structure Feels Metaphysically Hard

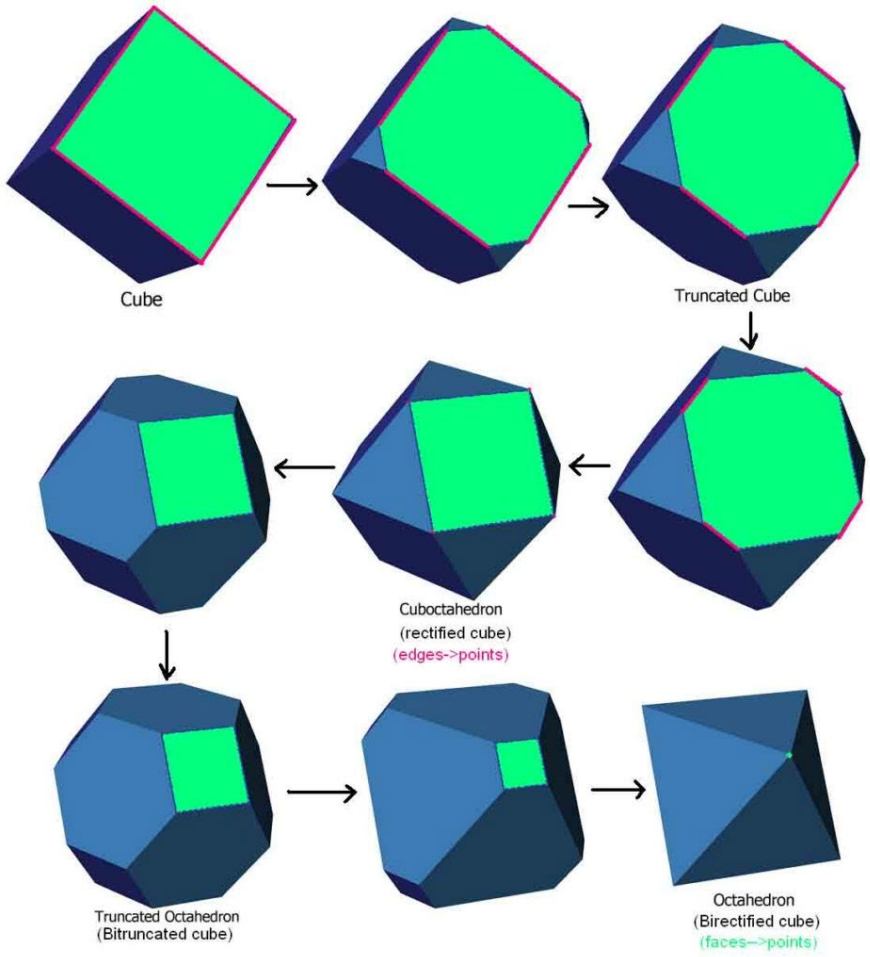

Figure 1. A "polyhedron movie" of chart change from cube to octahedron (via truncation and rectification). This sequence visualizes a continuous one-parameter deformation of the cube in which the *preferred discrete description* (the "chart" of faces/edges/vertices) changes at recognizable boundary events. Starting from the cube, progressive truncation rounds the vertex neighborhoods and introduces new faces, yielding the truncated cube. At the special midpoint (rectification), the cuboctahedron appears: the original *edges* have effectively become *vertices* (a refactorization of incidence). Continuing the deformation reverses which elements remain prominent: the original square faces shrink while triangular structure dominates, passing through the truncated-octahedron/bitruncated regime, and finally reaching the octahedron, where the cube's former faces have collapsed to points (faces→points).

Read as perception theory rather than solid geometry, the diagram illustrates the paper's claim: the "3-faces-at-a-vertex" versus "4-faces-at-a-vertex" difference is not best treated as an ontological snap, but as a chart transition inside a smoothly varying underlying structure, with brief critical zones where one incidence summary becomes ill-conditioned and another becomes stable.

We begin with a simple observation that becomes philosophically sharp the moment one tries to visualize it rather than merely state it. Some properties of shapes feel like "hard facts," not because they are logically deeper than others, but because the human perceptual system treats them as atomic. When a property is experienced as atomic, any change in that property is experienced as a rupture. The mind says: something snapped. Yet the mathematician says: there is a continuous family, there are limits, there are degenerations, there are dual descriptions, and nothing "mystical" has happened—only a shift of representation. The tension we investigate is precisely this: the existence of a formally continuous story alongside an intuition that insists the story must be discontinuous.

The cube–octahedron pair provides a particularly clean arena because it is not merely "two different solids." They are duals in the combinatorial sense: vertices of one correspond to faces of the other, and the shared number of edges functions like an invariant that tempts the imagination to search for a smooth bridge. But the vertex valence differs: in the cube, three faces meet at a vertex; in the octahedron, four faces meet at a vertex. That difference is not a matter of small deformation within a fixed combinatorial type. It is a discrete jump. If the visual system treats "faces meeting at a vertex" as a basic perceptual primitive, then any path from cube to octahedron must feel like a phase transition, not a morph. The discontinuity is not in the mathematics alone; it is in the metaphysics of perception.

Our project is to name that metaphysics and then imagine, with restraint, a different one.

1.1 The cube–octahedron duality as a perceptual stress test

Duality is a kind of promised reconciliation. It says: what you see as "a vertex" can be re-seen as "a face" in a systematic way; what you treat as "the meeting of faces" can be reinterpreted as "the adjacency of vertices," and the edges sit as mediators in both accounts. This is already a clue that the crispness of our usual ontology—faces, edges, vertices as irreducible things—may be a convenience rather than a necessity. Duality suggests that the same underlying incidence information can be packaged in two symmetric ways. In a purely formal setting, this is untroubling: one swaps roles of elements in a structure and obtains another structure with corresponding invariants.

But as a perceptual thought experiment, it becomes stressful. If the cube and octahedron are "the same story told with roles swapped," why does the mind fail to experience a smooth story

that passes between them? If edges match, why does the subjective impression remain: “these are categorically different objects”? The stress test works because it puts pressure on a tacit assumption: that perception is basically a faithful reading-off of discrete parts of reality. Duality undermines that assumption. It implies that there is no unique privileged decomposition into primitives, even in the apparently humble case of convex polyhedra. If no decomposition is privileged, then the mind’s insistence on one decomposition as “what’s really there” becomes visible as a choice, or even as a bias.

The cube–octahedron pair also adds a second layer of pressure: the mismatch between invariants and experience. The shared number of edges is an invariant that many people can hold in mind simultaneously with the dual correspondence. Yet the mismatch in vertex valence is experienced as a discontinuity that dominates the imagination. The invariant does not soothe the discontinuity. Instead, it sharpens it: “how can something stay the same in one respect and yet be forced to change in this other respect that feels so basic?” In other words, duality does not merely relate two solids; it exposes which features of the solids our perceptual system treats as metaphysically primary.

This makes the pair a diagnostic tool. A mind that can treat cube↔octahedron as an ordinary dual re-description is a mind that has already loosened its grip on object-first primitives. A mind that cannot help but feel the discontinuity is revealing something about its native representational grammar.

1.2 Why “3 faces vs 4 faces” resists smooth visualization in object-first vision

In object-first vision, the perceptual pipeline is dominated by segmentation: boundaries, edges, corners, and stable parts. This is not an arbitrary claim; it is an empirical fact about what human vision is good at. We are efficient at locking onto occlusion boundaries, at carving scenes into objects, at stabilizing identity under motion, and at compressing continuous sensory flow into discrete, reusable tokens: face, hand, chair, corner, rim, eye. Those tokens are not merely linguistic; they are perceptual. They feel immediate.

When “face,” “edge,” and “vertex” are treated as tokens of this kind—primitives in the mind’s internal scene graph—then incidence becomes a Boolean relation: attached or not attached. The count “three faces meet here” becomes part of the identity of the vertex, not a derivative statistic. To change that count is to change what the vertex is. The perceptual system therefore experiences a nonnegotiable discontinuity.

More precisely: valence (how many faces meet at a vertex) is a combinatorial invariant of the polyhedral cell decomposition. If you insist that the object remains a polyhedron with planar faces and a stable incidence structure, then valence cannot drift continuously. It can only change through an event where the decomposition changes: a face shrinks away, a new face is

created, an edge collapses, a vertex splits, or some other "critical" reconfiguration occurs. Even if the ambient geometry varies continuously, the label "valence = 3" does not vary continuously. In object-first perception, labels of this kind are treated as the bedrock of what is being seen. Therefore, continuous geometric variation that requires a relabeling is experienced as an ontological glitch: something that "should not happen smoothly."

This is why many otherwise reasonable bridges—truncation movies, smoothing families, norm-ball interpolations—often fail to address the specific discomfort you identified. They show a continuous change of shape, but they do not make the valence switch feel ordinary. The mind keeps returning to the discrete census: "At the corner, how many faces are there?" It keeps looking for the moment the fourth face appears, and when it appears, it feels like magic. The continuity of the surface is not the continuity the mind cares about, because the mind's ontology is not "surface as a field." It is "polyhedron as a set of faces." In that ontology, a new face is a new entity, and entity-creation is a rupture.

One can summarize the resistance like this. Object-first vision binds identity to parts. But the cube↔octahedron transition binds identity to something deeper than parts: to the possibility of re-partitioning the same incidence information under a duality. The mismatch is metaphysical, not computational. The visual system is not merely "bad at seeing it"; it is committed to a different notion of what counts as continuity.

1.3 The paper's wager: change the perceptual primitives, not the geometry

The wager of this paper is that the felt discontinuity is not a brute fact about reality; it is a fact about a particular representational choice. If we change what the mind takes as primitive—if we change its "native nouns"—then what looked like a phase transition can become an ordinary re-description. We are not proposing that the cube literally deforms into the octahedron without passing through any critical event in the combinatorial sense. We are proposing something subtler: that a mind may track continuity at a deeper level than the discrete census of faces at a vertex, and may treat the census as an emergent readout rather than a metaphysical axiom.

In the CT/HoTT-flavored hypothesis, the perceptual primitives are not parts but transformations and witnesses. The mind does not begin with faces; it begins with composable relations in the sensory stream. It does not begin with equality; it begins with identification paths and coherence across alternative parses. In such a mind, "3 faces meet" and "4 faces meet" are not competing claims about what is really there; they are two stable decompositions that become appropriate in different regimes of a continuous underlying state. The valence switch is then experienced not as creation ex nihilo, but as a chart change: the same underlying structure expressed in a new coordinate system when the old one becomes ill-conditioned.

This paper therefore does not aim to teach a reader to visualize a valence jump within the human visual cortex as currently constituted. Instead, it aims to describe a hypothetical perceptual metaphysics in which such jumps are no more shocking than a shift in perspective, a change of lighting, or the recognition that two different sketches depict the same object. The cube–octahedron example is our anchor because it is small, rigid, and clean. But the purpose is broader: to explore what it would mean for perception—especially of faces and ordinary objects—to be equivalence-first rather than object-first, and how "identity as path" would alter what counts as a smooth, ordinary experience.

2. Two Kinds of Seeing: Object-First Versus Relation-First

To speak about a "visual cortex wired with CT/HoTT" without drifting into empty metaphor, we need a clean contrast class. The contrast is not "mathematical people versus non-mathematical people," and it is not "humans versus aliens" in any crude sense. The contrast is between two representational commitments that a mind can, in principle, embody. Each commitment answers a quiet question that perception must answer continuously: what is basic, and what is derived?

In an object-first metaphysics, the basic units are things and their boundaries. In a relation-first metaphysics, the basic units are constraints and the compatibility of multiple partial descriptions. Both metaphysics can recover the same everyday competence—navigating the world, recognizing faces, reaching for a cup—but they do so with different primitives. The cube–octahedron stress test is useful precisely because it forces these primitives into view: it exposes what each metaphysics treats as "obviously real."

2.1 Object-first vision: edges, parts, segmentation, stable labels

Object-first vision is the human default, and it is so successful that it becomes invisible. The world arrives as a continuous sensory flow, yet within a fraction of a second we experience discrete entities: a face, a chair, a hand, a doorway, a shadow on the floor. This discreteness is not merely conceptual; it is perceptual. The mind assigns boundaries, and the boundary assignments stabilize into objects. Once an object is stabilized, perception becomes economical: we can track it across time, re-identify it under viewpoint changes, and predict its behavior.

Several features characterize object-first vision.

First, it privileges edges and discontinuities. An edge is where something changes abruptly: brightness, color, depth, texture, motion. Edges are highly informative, and object-first systems treat them as fundamental evidence for "where one thing ends and another begins." Corners

become special because they are stable under viewpoint changes; they are anchors for recognition and manipulation.

Second, it treats parts as real. A face is not merely "a surface"; it is eyes, nose, mouth, jawline. A car is not merely "a rigid body"; it is wheels, headlights, windshield. Part structure becomes a major axis of recognition and prediction. Importantly, part structure is discrete: either the wheel is there or it is not, either the mouth corner is upturned or it is not. The world is thus compressed into a set of named, reusable components.

Third, it relies on stable labels. Once segmentation and parts are in place, the system can attach labels that persist: this is the same person, that is the same cup, this is still the same dog. The labels are not always conscious, but they are functionally real. They support rapid inference and social life. They also encourage a metaphysical habit: to treat the labeled objects as the "real units" of reality.

Object-first vision is not naive; it includes sophisticated constancy mechanisms. It does not collapse when lighting changes, or when a face rotates. But it achieves constancy largely by preserving a stable object token through transformations, rather than by treating the transformations themselves as primary. The transformations are handled, but they are handled as "nuisances" that must be factored out to recover the stable object and its parts.

This is why object-first vision struggles with certain phase-change stories. If a polyhedron is represented as a set of faces meeting along edges at vertices, then a change in how many faces meet at a vertex is a change in the object's part-structure. And if part-structure is metaphysically primary, then the change feels like a rupture.

2.2 Relation-first vision: constraints, adjacency tendencies, coherence fields

Relation-first vision begins from a different assumption: the world is not first given as discrete things, but as a web of regularities and constraints, and "things" are stable patterns within that web. In this metaphysics, edges and parts are not denied; they are recovered as convenient summaries when they are warranted. But the primitive currency is not "object with parts." It is "constraint satisfaction over a field of possibilities."

Several features characterize relation-first vision.

First, it privileges constraints over boundaries. A constraint is a lawful compatibility condition: surfaces cannot interpenetrate; rigid parts move together; illumination changes smoothly across surfaces; facial muscles deform in correlated ways; two eyes tend to be symmetric about an axis and co-move with the head. A relation-first system represents these compatibilities as primary. It "feels" the world as something that must hang together, rather than as a set of segments to be labeled.

Second, it treats adjacency as a tendency rather than a switch. In object-first vision, adjacency is often Boolean: either these regions meet or they do not. In relation-first vision, adjacency can be graded: regions can be near-contact, almost-coincident, weakly coupled, strongly coupled. This matters for the cube–octahedron problem. A new face "appearing" is only metaphysically shocking if the mind believes faces are discrete entities that blink into existence. A relation-first mind can treat "facehood" as a regime of a continuous underlying state: a planar tendency becoming dominant, a constraint becoming tight, a coupling becoming strong enough to count.

Third, it maintains multiple partial parses and measures their coherence. Where object-first vision often commits early—this boundary is real, this part is an eye—relation-first vision can remain agnostic longer. It can hold competing decompositions and evaluate them by coherence: which decomposition yields a smoother, more consistent explanation of the evolving sensory stream? This naturally aligns with a HoTT-flavored idea: identity is not a single label; it is a family of identifications supported by witnesses of compatibility.

Fourth, it experiences "critical events" as ordinary. In a world of constraints, it is normal for regimes to shift when a parameter crosses a threshold: contact becomes separation, slip becomes stick, resonance locks in, a shadow becomes an occluder, a specular highlight appears. Relation-first perception treats such thresholds as part of the normal texture of seeing, not as metaphysical alarms. This is the experiential groundwork needed to treat a combinatorial transition—like a vertex valence change—as a kind of regime shift rather than a rupture of being.

In a CT/HoTT-wired hypothesis, relation-first perception would be especially attentive to composition and coherence. It would experience not merely "what is there," but "how the evidence composes," "how different explanatory paths agree," and "where coherence fails." In that setting, the cube and octahedron are not two unrelated objects. They are two stable presentations of an incidence structure, and a transition between them is a reorganization of constraints that remains continuous at the level of coherence.

2.3 What "ordinary" means in each metaphysics of perception

The word "ordinary" can mislead. It does not mean "simple." It means "not metaphysically alarming." Something is ordinary when it does not trigger the sense that reality has broken. In this paper, "ordinary" is the name of what a perceptual metaphysics treats as continuous, expected, and non-paradoxical.

For object-first vision, ordinary means: stable objects persist through time, parts remain parts, boundaries remain boundaries, and identity labels remain anchored. Ordinary change is change that preserves the object-token: a face turns, a hand moves, a cube rotates. Extraordinary change is change that forces a re-tokenization: when the object appears to change its part-

count, when boundaries flip, when something that was “one object” becomes “two,” or when the system must revise what it thought the object was. In this metaphysics, discontinuities in part structure are felt as unusually deep, because part structure is close to ontological bedrock.

For relation-first vision, ordinary means: coherence persists even when the preferred decomposition changes. Ordinary change can include re-factorizations: a new segmentation becomes more coherent than the old one, a different “chart” becomes the better description, a new constraint becomes active, and yet the underlying continuity of the scene remains felt. Extraordinary change is not “a part appeared,” but “coherence failed”: the system cannot find a compatible set of constraints that explains the stream. In other words, the alarm condition is not discrete re-description; it is inconsistency that cannot be repaired by changing the description.

This contrast explains why the same physical scenario can be experienced very differently. In object-first seeing, the cube↔octahedron story is not primarily about continuous deformation; it is about a forbidden alteration in the census of parts. In relation-first seeing, the same story is about a continuous shift in which constraints are tight and which decomposition is most coherent. The valence switch is then no more metaphysically shocking than the moment a rotating object’s silhouette changes: the preferred representation changes, but the felt continuity of the world does not.

The point of this section is not to declare one metaphysics “correct.” Both are plausible strategies for finite agents in a noisy world. The point is to isolate the representational commitments that make a vertex valence change feel like a rupture, and to identify what commitments would make it feel ordinary. Once those commitments are clear, the later sections can describe how a CT/HoTT-wired cortex would implement relation-first seeing, and what that would imply for the perception of ordinary objects such as human faces.

3. A Minimal Technical Backbone (Kept Intuitive, Not Formalist)

This paper is not an attempt to “explain category theory” or “explain HoTT” in the usual didactic sense. We need only a minimal backbone—just enough structure to make our metaphysical claims precise enough to be testable in thought experiments. The guiding principle is: if a mathematical notion is invoked, it must correspond to a plausible perceptual operation. Otherwise it is decoration.

The key move is to treat perception not as the production of static internal pictures, but as a stream of transformations that can be composed, compared, and validated for coherence. Once perception is understood as something that composes, the language of categories becomes

natural. Once “sameness” is understood as something witnessed by continuous transformations rather than asserted by a label, the core intuition of HoTT becomes natural.

3.1 Categories as “composable perceptions” rather than static images

A category is often introduced as “objects and arrows.” In a perception setting, that slogan can be misleading because it tempts the reader to think “objects = things in the world.” Instead, the useful interpretation is operational: a category is a way of organizing processes that can be chained.

In ordinary seeing, we perform a great many micro-operations: we track a contour over time, we integrate motion signals, we stabilize an object under viewpoint change, we infer a 3D pose from 2D cues, we correct for illumination, we bind features into a single identity token, we discard transient noise. These are not single acts; they are composable operations. You do not “recognize a face” in one step. You do a sequence of transformations on the incoming stream until the stream is in a form where recognition becomes stable.

This is the perception-friendly reading of a category:

- The “objects” are not physical objects. They are internal representational states: raw sensory snapshots, intermediate feature maps, candidate segmentations, hypothesized poses, stabilized identity states, and so on.
- The “morphisms” (arrows) are the transformations that carry one representational state to another: filtering, grouping, tracking, pose-estimation, normalization, attention shifts, and inference steps.

Composition is the point. If you can go from state A to state B and from state B to state C, you can go from A to C by composing the operations. That is exactly how perception behaves. It is a pipeline, but not a rigid one. It can branch, compare, and revise. It is also not necessarily linear in time; perception frequently revisits earlier hypotheses in light of later evidence.

Two additional categorical ideas matter here, even at an intuitive level.

First, multiple paths can lead to “the same place.” You can infer a face identity by tracking the whole head, or by using eye-region features, or by integrating voice cues with visual cues, or by using gait and context. These are different routes through representational space. A category gives us a language for talking about the existence of multiple routes and about when they agree.

Second, there is a difference between “arriving at the same label” and “arriving at the same structure.” Object-first perception often confuses these: it wants a single stable label, and it treats label agreement as the victory condition. A category-aware perception would treat

agreement across different compositional routes as evidence of something deeper: not merely that the system declared the same label twice, but that the system found a coherent structure that is stable under different transformations.

Thus, the claim "a CT-wired cortex sees compositional structure" can be rendered non-mystically: it means the cortex treats perception as a network of composable transformations whose mutual compatibility is perceptually salient.

3.2 Functorial perception: structure-preserving mapping from sensation to shape

A functor is, in essence, a map between categories that preserves the compositional structure. Again, the usual mathematical phrasing can distract. The perceptual translation is straightforward: a functor is a disciplined way of translating one kind of description into another without breaking the relations that matter.

Let us name two domains:

- The domain of sensation: time-indexed sensory data, with its own native relations (temporal adjacency, correlation, continuity of motion, smoothness of illumination, consistency of optical flow).
- The domain of shape: objects, surfaces, incidences, poses, and identities, with their own native relations (adjacency of parts, rigidity constraints, continuity of deformation, symmetries, persistence through time).

If perception is functorial, then the mapping from sensation to shape does not simply spit out a best guess at each time slice. It carries over the structure of the sensory relations into structural relations in the inferred world. Roughly: "nearby in sensation time" becomes "nearby in inferred world-time"; "consistent motion cues" becomes "same moving object"; "co-moving features" becomes "rigid part"; "stable invariants under viewpoint change" becomes "identity persistence."

This sounds obvious, but it has a critical philosophical consequence. It means the cortex is not merely doing classification; it is doing structure translation. The output is not just a label ("face") but a structured object in a world-model, and the structure in the output is constrained by the structure in the input.

This is where the cube–octahedron problem becomes legible. The object-first mind wants the output to be a fixed decomposition: a polyhedron with faces meeting at vertices, and "valence" is then a literal property of the output object. A functorial, relation-first mind can treat the output not as a fixed decomposition but as a structure class: a family of decompositions connected by transformations that preserve the relations that the sensory evidence supports.

In that setting, the mind does not need to decide prematurely whether the local configuration is best summarized as “three faces meet” or “four faces meet.” It can preserve a more primary relational structure—say, an adjacency network or constraint pattern—and then allow multiple decompositions to be valid “charts” on that structure depending on regime. The functoriality claim is: the cortex preserves composition and coherence when moving from sensation to world-model, and therefore it naturally treats “re-description” as something that can be continuous even when discrete summaries change.

There is a second reason functorial language matters here. It clarifies what it would mean for perception to be “CT-like” without claiming that the brain literally manipulates formal categories. The claim is about invariants: the brain might be selecting representations that keep important compositions valid and keep multiple explanatory routes aligned. That is a functional hypothesis, not a claim about neurons spelling out commutative diagrams.

3.3 HoTT’s core move: identity as paths, not equality as a switch

The key HoTT intuition we need is not its full technical machinery. It is one metaphysical relocation: identity is not treated as an external stamp, but as something internal to the space of appearances.

In everyday cognition, we often treat identity as a switch:

- Same person or different person.
- Same object or different object.
- Same shape or different shape.

This is cognitively efficient. It supports fast action. But it also makes certain transitions feel metaphysically hard, because a switch has no internal structure. When the switch flips, we experience a jump.

HoTT replaces the switch with a path. “A is the same as B” is not merely a proposition; it is witnessed by a continuous transformation—an identification—and there can be many such identifications, related to each other by higher transformations. In our perceptual translation, this becomes:

- “This face now” and “that face a moment later” are the same person because there exists a continuous transport through appearance space that carries one to the other while respecting constraints (rigidity, plausible muscle deformation, consistent lighting changes, continuity of motion).

- “This local polyhedral neighborhood described one way” and “the same neighborhood described another way” are the same structure because there exists a transformation that witnesses their equivalence, even if the discrete census of parts differs.

Once identity is path-like, several consequences follow that matter for our thesis.

First, identification has thickness. It is not all-or-nothing. A path can be short or long, easy or hard, stable or fragile. A mind can therefore experience “sameness” with a gradient: strongly identified, weakly identified, identified but with strain, identified but only via a complex route. This supplies the phenomenological room for “a fourth face is emerging” without implying that a new entity popped into existence.

Second, ambiguity can be stable. If identity is a set of possible paths rather than a switch, the mind can hold multiple candidate identifications and wait for coherence to decide. This aligns with relation-first vision and makes chart changes ordinary.

Third, failure becomes meaningful. In a switch-based system, when recognition fails, one often gets only a binary outcome: not recognized. In a path-based system, failure has a geometry: the attempted transport breaks at a particular place, the constraints become incompatible, the coherence witness cannot be constructed. This is what would make incoherent stimuli (including certain synthetic or manipulated signals) perceptually salient as “does not lift,” “does not cohere,” “cannot be transported,” rather than merely “looks weird.”

Finally, the cube–octahedron valence issue can be reframed. The reason “3 vs 4” feels hard is that it is treated as a switch in the object’s description. A HoTT-flavored perceptual metaphysics would instead treat the underlying continuity as primary and the discrete census as a chart-level artifact. The mind would not deny that the chart-level artifact changes. It would simply not treat that change as metaphysically shocking, because it experiences the continuity through the existence of a path of identifications that remains intact across the regime change.

This is the minimal backbone we will carry forward: perception as composable transformations, mapping that preserves structure, and identity as witnessed transport rather than asserted equality. With these three ideas in place, we can discuss polyhedral phase changes and ordinary face perception in a single unified vocabulary: coherence, chart choice, transport, and the ordinary life of paths.

4. Polyhedral Phase Change Reframed

The cube–octahedron episode is a small laboratory for a large question: what counts as a “smooth change” for a mind? In ordinary object-first perception, smooth change is defined as change that preserves the same discrete parts and their incidence relations. But polyhedral

duality, and especially the valence mismatch (3 faces at a cube vertex versus 4 at an octahedron vertex), forces us to separate two notions of continuity that human intuition tends to fuse:

1. Continuity of an underlying geometric or constraint state.
2. Continuity of a chosen discrete description (faces/edges/vertices with fixed incidence).

The reframing in this section is: treat (1) as primary and (2) as a chart-level readout. Then a "phase change" is no longer an ontological shock; it is a routine shift in which chart best summarizes the same underlying state.

4.1 Valence as a derived census, not a primitive fact

Valence is the count of how many faces meet at a vertex. In the combinatorial description of a polyhedron, that count is crisp and nonnegotiable. It is tempting to treat it as metaphysically basic: if the valence changes, surely the object has become a different object.

The relation-first/CT/HoTT wager is to demote valence from bedrock to readout.

To do that, we must say what replaces it. The replacement is not "vague blur." It is a richer underlying description that can remain continuous while discrete counts change. Several candidates are plausible; we only need the idea, not a single canonical choice:

- A constraint description: which planar tendencies are strong, which adjacency relations are tightening, which dihedral-angle pressures are increasing or decreasing.
- An implicit-surface description: a shape specified by an underlying field, where "faces" are emergent planar regimes rather than primitive polygons.
- A skeletal or medial description: a branching structure that captures how the shape organizes space, which can undergo bifurcations while remaining continuous in a parameter.

In each case, "valence" becomes something you compute after the fact, by applying a discretization rule: pick planar patches, threshold them into faces, identify their intersections, and count. The key metaphysical shift is: the count is not what the mind is primarily tracking. The mind is tracking the continuous structure underneath, and the count is a late census.

This aligns with ordinary experience in less mathematical contexts. Consider a shadow boundary on a face. At one moment you might parse it as "a line," later as "a curve," later as "a patch of darkness," depending on lighting and viewpoint. The underlying light field changes smoothly; the discrete descriptions you use to summarize it can change. The point is not that the world is ambiguous; it is that the mind's discrete tokens are summaries chosen for utility. A

CT/HoTT-wired mind would apply the same humility to polyhedral parts: “faces” are convenient summaries, not inviolable atoms.

Once valence is treated as a census, the question “how can 3 become 4 smoothly?” becomes: “how can the underlying structure shift smoothly so that the best census changes from 3 to 4?” That is no longer paradoxical; it is the normal logic of thresholded readouts.

4.2 Stratified spaces: smooth regions plus boundary events as normal loci

The next step is to normalize the idea that continuity and discreteness coexist. The appropriate mental model is not a single smooth manifold where everything varies differentiably. The appropriate model is a stratified space: a space built from layers (strata) where each layer is smooth, but different layers meet along boundaries where qualitative descriptions change.

This is not an exotic idea in lived reality. Many ordinary phenomena are stratified:

- Solid versus slipping contact (stick-slip friction): a smooth change in applied force can trigger a discrete regime shift.
- Phase change in matter: temperature changes smoothly, state changes discretely.
- Resonance lock-in: a parameter varies smoothly, synchronization appears abruptly.

In all these cases, the “event” is not a violation of continuity; it is the crossing of a boundary between regimes. A mind that is comfortable in stratified spaces does not treat boundary crossings as metaphysical alarms. They are expected loci where the description changes.

Apply this to polyhedra. The “space of shapes” can be thought of as containing regions where a given combinatorial structure is stable—where a particular face/edge/vertex incidence pattern is robust under small deformations. Those regions are separated by boundaries where some feature degenerates: an edge shrinks to zero, a face area collapses, a dihedral angle hits a limit, a vertex becomes coplanar with something it was not. At those boundaries, the best combinatorial description changes.

Object-first vision treats the boundary as “where the object stops being itself.” A CT/HoTT-flavored relation-first perception treats the boundary as a normal locus in the space of presentations: a place where chart A becomes ill-conditioned and chart B becomes the coherent replacement. Continuity is tracked not by insisting the same chart works everywhere, but by tracking a coherence-preserving identification that remains intact across chart changes.

This is exactly the perceptual reframe needed to make valence switches feel ordinary. A valence switch is a boundary event: not “reality snapping,” but “we crossed into a region where a different census is appropriate.”

4.3 Degeneration moments: where a new incidence becomes fully legible

If we want the reframing to be more than a slogan, we need to say what happens at the moment of transition—what we previously called the "degenerate" moment. Degeneration is the stage where a discrete summary becomes unstable because one of its components is losing measure: a face is shrinking, an edge is collapsing, an intersection is becoming ambiguous.

In object-first metaphysics, this is experienced as failure. A face that becomes arbitrarily small looks like a glitch: either it is a face or it is not. A system forced to label it must eventually flip the label, producing the felt snap.

In relation-first metaphysics, degeneration is experienced as legibility changing. The system tracks not "is face present?" but "how strongly is this planar regime supported?" or "how stable is this adjacency under small perturbations?" As the feature shrinks, its support weakens; its role in the coherence network decreases. It is not suddenly annihilated; it fades in structural relevance. Conversely, a new incidence—what will become "the fourth face at the vertex"—can be treated as becoming increasingly supported before it crosses whatever threshold makes it count as a face in the census.

This is the graded-incidence idea in action. It is not that the world is fuzzy; it is that the mind has a continuous measure of how warranted a discrete token is. Humans do this in many domains already. We can feel a perception "solidify." We can feel an interpretation "become obvious." We can feel a recognition "snap in." That snap is often not a physical discontinuity; it is a threshold in internal confidence.

A CT/HoTT-wired mind would treat degeneration moments as ordinary precisely because it does not equate "a discrete token exists" with "reality exists." It equates reality with the coherence of constraints. A new incidence becomes "fully legible" when the coherence witness that supports it becomes stable and reusable across compositional paths of inference.

Here is the important metaphysical claim: in such a mind, the "moment of creation" is not a primitive metaphysical event. It is a representational event: the moment a new chart becomes stable enough to be the default. The underlying structure never needed a sudden act of creation; it needed a crossing of a boundary where one summary becomes less coherent than another.

4.4 Duality as binocular perception: primal and dual co-present

So far we have reframed phase change as chart change in a stratified space. Duality adds a second stabilization mechanism: it gives the mind two coupled descriptions that can mutually confirm coherence.

In the cube–octahedron pair, the primal description emphasizes faces meeting at vertices; the dual description swaps those roles. The same incidence information is being packaged differently. If a mind can hold both packages as co-present, then it does not have to cling to one set of primitives. It can treat “vertex facts” and “face facts” as two views of a single underlying structure.

This is where “binocular” is more than metaphor. Human binocular vision fuses two slightly different images into depth. The fusion is not a logical inference we perform consciously; it is a native percept. The suggestion here is analogous: a CT/HoTT-wired perception might fuse primal and dual descriptions into a single coherent percept of incidence. The mind would not need to pick whether vertices or faces are primary; it would experience them as dual projections of one structure.

What would this do to the valence problem? It would de-privilege the vertex census. In primal mode, the mind might notice “three faces meet here.” In dual mode, the same local region is understood through a corresponding face in the dual. The dual face has its own invariants (e.g., number of edges bounding it). As the system approaches a boundary event, the primal census may become unstable, but the dual description may remain stable, or vice versa. The mind’s felt continuity can be anchored by whichever view remains coherent, while the other view is re-charted.

In this way, duality acts like perceptual redundancy. It provides a second coordinate system that prevents a local singularity in one chart from being experienced as a metaphysical rupture. The mind experiences: “one description is becoming ill-conditioned; the other still carries the structure; I can transport the identification across; now I return with a new, stable primal chart.” That is exactly the kind of ordinary experience we want to model: a re-description that feels like turning an object in the hand, not like reality tearing.

This binocular-duality hypothesis also foreshadows why face perception is a good next test case. Human face recognition already uses multiple redundant cues: geometry, texture, motion, symmetry, expression dynamics, context. A CT/HoTT-wired mind would treat such redundancy not merely as “more features,” but as multiple compositional routes whose mutual agreement constitutes coherence. Duality in polyhedra is then a toy version of what ordinary perception might already do in richer form: hold multiple presentations of the same underlying identity and experience continuity as the existence of stable transports between them.

In summary, reframing polyhedral phase change requires four shifts. Valence becomes census, not essence. The space of shapes is stratified, so boundary events are normal. Degenerations are moments of changing legibility, not creation ex nihilo. Duality supplies a second co-present

chart, allowing continuity to be experienced as coherence-preserving transport rather than as preservation of a single discrete decomposition.

5. What a CT/HoTT-Wired Visual Cortex Would Represent

To say "a cortex wired with CT/HoTT" cannot mean that neurons are literally proving theorems. The claim has to be representational: the primitive formats in which the system stores and updates percepts are organized like categorical and homotopical ideas. That is, the system is built to (i) preserve compositional structure, (ii) treat identity as transport, and (iii) register coherence not as an afterthought but as a primary perceptual signal. The most important consequence is that what humans experience as "objects" would be, for this mind, derived compressions of a richer internal state.

This section describes that internal state as four interlocking representational commitments. They are presented as metaphysical "what it is like" claims, but each is also a computational hypothesis: a way a perceiver could remain robust in noise while avoiding premature hard discretization.

5.1 Percepts as types with families of presentations (charts)

In an object-first cortex, the basic representational unit is an object token: a bounding region in the visual field with a label and a set of features. That token is maintained through time by tracking and updating. The token is the anchor of identity: if the token persists, the object "is the same."

In a CT/HoTT-wired cortex, the anchor would be different. The basic unit would be closer to a type: not a single instance, but a structured space of allowable instances together with the ways they can vary. A "face," a "hand," a "cube," a "person's head" would not be stored as a fixed picture but as a family of presentations—multiple coordinate systems that can each represent the same underlying entity in different regimes.

A chart, in this perceptual sense, is a particular presentation format that makes certain aspects easy to represent. Examples in ordinary perception already exist:

- A face can be represented by a sparse landmark configuration (eyes, nose tip, mouth corners).
- The same face can be represented by a texture map aligned to a canonical template.
- It can be represented by a 3D head model with pose parameters.
- It can be represented by a motion signature across time.

Humans often move between such representations implicitly, but we typically experience only the stabilized token: “that person.” A CT/HoTT-wired system would experience the family itself as primary. It would not regard a single chart as “the real face.” It would regard “Face” as a type with multiple charts, each chart valid in a region of conditions (lighting, occlusion, distance, angle, expression).

This immediately reframes the polyhedral problem. “Cube” and “octahedron” become not merely two objects but two charts on an underlying incidence type. The mind can remain continuous by remaining in the type while changing charts. The discrete census (faces meeting at a vertex) is a chart-dependent readout. If one chart becomes ill-conditioned near a boundary event, the mind does not panic; it transitions to a chart that remains stable.

Phenomenologically, this would feel like a kind of “representational depth.” Where we see one stable picture, this mind sees a stable entity plus the latent availability of alternative views, each one already partially present as a way the entity could be re-described.

5.2 Attachments as graded participation (soft incidence)

The next commitment concerns relations. Object-first perception tends to encode relations as discrete: adjacent/not, attached/not, part-of/not. This is not because the world is discrete, but because action often demands discrete decisions. If you are reaching for a cup, you need a crisp boundary: grasp here, not there.

A CT/HoTT-wired cortex can still act crisply, but it need not represent crisply at the base layer. It can treat attachments as graded participation. Instead of “this face meets this vertex,” it maintains something like “this planar regime participates in this junction with strength s,” where s varies continuously.

The metaphysical upshot is that “incidence” becomes a living, analog quantity rather than a Boolean predicate. This is exactly what is needed to make valence changes ordinary. The fourth face does not spring into existence; it goes from near-zero participation to high participation. The old three can simultaneously weaken or reconfigure in participation. At some point, a thresholded summary will count four rather than three, but the felt experience is continuous because the underlying participations have been drifting continuously all along.

This notion is not as alien as it sounds. Humans already represent many relations as graded, even if language forces a discrete report. We perceive “almost touching,” “barely separated,” “strongly grouped,” “weakly grouped,” “belonging together,” “accidentally aligned.” In social perception we do this constantly: “they are close friends,” “they are acquaintances,” “they are drifting apart.” The difference is that we rarely treat such gradedness as ontological for geometry. We revert to hard counts: how many faces, how many vertices, how many edges.

A soft-incidence cortex would treat geometry more like social grouping. A “vertex” is a junction where multiple constraints converge, each with a weight. A “face” is not a polygonal primitive; it is a stable planar tendency that participates strongly over a region. Edges are loci of competing tendencies. In such a world, discrete polyhedral descriptions are just one way to compress the continuous participation structure when the situation warrants it.

5.3 Higher morphisms as perceptual continuity of re-factorizations

The third commitment is the one that most strongly distinguishes a CT/HoTT-wired cortex from a merely “fuzzy” one. The point is not to blur everything. The point is to represent not only states and transformations, but transformations between transformations—what HoTT and higher category language call higher morphisms.

Translated into perception, a higher morphism is a witness that two different parsing routes, two different decompositions, or two different transformation sequences are not merely both available, but are compatible and continuously related. This is what makes re-factorizations feel continuous rather than like arbitrary switches.

Consider an ordinary example. You can recognize a person from a frontal face image, or you can recognize them from a side profile and gait and context. Those are different routes. A higher-morphism percept would register not just “both routes say it is the same person,” but “the routes can be reconciled”—there exists a continuous translation between the internal representations produced by each route.

Now map that to the cube–octahedron scenario. A re-factorization would be a change in the local decomposition: the way the surface is partitioned into faces, edges, vertices. Object-first vision tends to treat this as a re-labeling event: face count changed, therefore something broke. A higher-morphism cortex would instead carry an explicit witness that the old decomposition and the new decomposition are two presentations of the same underlying structure, connected by a continuous transformation in representation space.

This is the perceptual analog of “chart transition functions” in geometry, but with an important psychological twist: the system experiences the existence of the transition as a kind of felt continuity. Where we experience a snap, it experiences a glide accompanied by an internal guarantee: not “trust me,” but “I can see the transport.”

Higher morphisms therefore function as a stabilizer against ontological panic. They allow the system to move between decompositions while keeping identity intact at a deeper level. The discrete pieces can change because the identity is not attached to the pieces; it is attached to the witnessed equivalence between presentations.

If one wanted to say this in a single metaphysical sentence: this mind experiences "sameness" not as sameness of parts, but as sameness of process—sameness of the web of transformations that can carry one presentation into another.

5.4 Coherence as a felt invariant: agreement across multiple parsing routes

Finally, we come to the perceptual invariant that replaces "part count" as the primary anchor: coherence. In this context, coherence means that different compositional routes through perception—different sequences of operations, different cues, different charts—arrive at mutually compatible representations.

In ordinary human cognition, coherence is largely hidden behind confidence. We say "I'm sure it's him," or "it looks wrong," but we typically do not experience the internal structure of why it is sure or wrong. A CT/HoTT-wired cortex would experience coherence more directly, as a kind of perceptual qualia: the scene hangs together, or it does not. This hanging-together is not a mood; it is an invariant that the system can test by attempting transports and comparing results.

Coherence becomes the mind's "ordinary." If coherence holds, then chart changes, re-factorizations, and discrete census flips are not metaphysically alarming. They are just different summaries becoming optimal in different regions. If coherence fails—if no transport can reconcile the parses, if the routes disagree irreparably—then the system experiences an alarm. The alarm is not "a new face appeared"; it is "the world model cannot be made consistent."

This is where the earlier talk of deepfakes and uncanny artifacts becomes relevant, but we need not rely on that domain. Coherence is a general organizing principle. A real human face has constraints: muscles deform with anatomical plausibility, shading changes with geometry, gaze dynamics have a characteristic coupling between eyelids and eyeballs, microexpressions have temporal structure. A coherence-first cortex would detect violations not as "it looks weird" but as "the constraints cannot be jointly satisfied."

Returning to polyhedra, coherence plays the role of the deeper continuity that survives the valence switch. The mind can accept that the discrete face decomposition has changed because the underlying incidence type is still coherently transportable across presentations. The same edges, the same adjacency tendencies, the same constraint satisfaction pattern persists, even as the chosen census changes.

In summary, a CT/HoTT-wired visual cortex would represent percepts as types with multiple charts, relations as graded participations, re-factorizations as connected by higher-witnessed continuities, and "ordinary reality" as the felt coherence of multiple parsing routes. This is the representational package that makes discrete combinatorial changes psychologically continuous

without denying that combinatorial descriptions are discrete. It relocates continuity from the census to the witnesses.

6. Phenomenology: What It Would Feel Like to "See" the 3→4 Switch Smoothly

If the earlier sections describe a representational architecture, this section describes what that architecture would feel like from the inside. The aim is not to romanticize alien cognition, but to render the difference concrete enough that the reader can compare it to their own perception. The core proposal is that the CT/HoTT-wired mind would not "see more faces." It would see a different kind of continuity. The ordinary unit of continuity would not be "the same set of parts persists," but "the same structure remains transportable across re-descriptions." In that world, a valence change is not a rupture; it is a change of the best local vocabulary.

We describe four phenomenological signatures: chart change qualia, lift-succeeds qualia, critical-event qualia, and the contrast case explaining why humans feel a snap.

6.1 The "chart change" qualia: re-description without ontological rupture

Imagine watching a shape whose local geometry is evolving in a way that, under a certain discretization, will change a count: three incident faces will become four. For a human object-first perceiver, the relevant question becomes: "When does the fourth face appear?" The attention is pulled toward the census. For the CT/HoTT-wired perceiver, the relevant question is different: "When does the current description stop being the most coherent chart?"

The subjective feeling of a chart change is not "a new thing popped into being." It is closer to the experience of switching interpretive lenses without losing the object. A mild human analog exists: when you look at a Necker cube and the perceived orientation flips, you do not experience a new cube being created. You experience a re-interpretation. The CT/HoTT mind would experience the 3→4 event in that spirit, but with an important upgrade: the flip is not arbitrary. It is guided by increasing ill-conditioning of the old chart and increasing stability of the new one.

So the qualia is something like:

- "The old description is getting strained."
- "It still works, but it's becoming awkward."
- "Now a cleaner description becomes available."
- "I slide into it, and everything becomes simpler again."

The decisive point is that the object retains its identity throughout. What changes is not being, but coordinate choice. The mind's sense of reality is anchored in the underlying type, not in the particular face/vertex decomposition that happens to be convenient at one moment.

If you wanted a tactile analogy: it is like rotating a complex tool in your hand. At some angles it is easiest to describe it by one set of surfaces; at other angles, different surfaces become salient. The tool never "changes." Your best description does.

6.2 The "lift succeeds" qualia: smooth transport of local structure

The second signature is subtler and more central. In our proposed metaphysics, "sameness" is felt as transportability: if a local structure can be carried forward through changes without breaking coherence, the mind experiences continuity. This is what we called "lift succeeds."

To understand the qualia, imagine that at each moment the mind has a local description of the shape near a junction (a would-be vertex). It also has an expectation of how that local description should evolve as parameters drift. The mind attempts to lift the local description along the changing conditions, meaning it tries to maintain the identification between "this local structure now" and "this local structure a moment later."

When the lift succeeds, the mind feels something like:

- "Yes, I can follow it."
- "The structure remains the same, even as its presentation changes."
- "I do not need to re-decide what I am seeing; it carries itself forward."

This feels less like counting and more like tracking a melody across a key change. The notes differ, but the relational structure is preserved and can be followed. The preservation is not merely inferred; it is experienced as a kind of smoothness.

At the moment the census changes from 3 to 4, the lift does not fail. The local structure has been continuously carried; the chart has changed, but the transport never broke. The mind therefore does not register a discontinuity. The "event" is not an appearance of a new face; it is a re-factorization that the mind can still lift across. The sense of continuity is anchored in the successful lift.

For humans, a weak analog occurs in skilled perception tasks where the brain tracks something through a transformation that novices experience as a jump: for example, musicians tracking harmonic function through modulations, or athletes tracking a ball through occlusion. In those domains, continuity is felt as "I can keep the thread." The CT/HoTT mind generalizes that to the very ontology of shape.

6.3 The “critical event” qualia: boundary crossing as ordinary perception

Now we come to the “phase change” itself. In a stratified world-model, boundary crossings are inevitable. The question is whether they feel like breakdowns or like ordinary transitions.

For the CT/HoTT mind, a critical event has a recognizable phenomenological character:

- The current chart becomes progressively ill-conditioned (increasing strain).
- A small set of degrees of freedom becomes dominant (the system’s attention narrows).
- Competing charts become closer in coherence score (a sense of imminent re-description).
- Then the system crosses a boundary and one chart becomes clearly superior again (release).

Subjectively, this might feel like passing through a narrow pass between two stable valleys. There is a brief region of heightened sensitivity where small perturbations could change which chart is chosen, but the mind does not experience this as unreality. It experiences it as a familiar kind of transition zone.

This matters because it makes the “critical event” a normal kind of perceptual episode. It is not the moment the world reconfigures itself; it is the moment the observer’s most efficient description changes. The mind expects such zones in the same way we expect the horizon line to shift when we stand up, or the apparent shape of a distant object to change when we move closer.

The key is that the mind’s metaphysics has been trained to treat boundaries as places where the world is most informative. A boundary is not a failure of smoothness; it is a place where the system learns which invariants are truly stable, because those invariants survive the passage.

In the cube–octahedron case, the mind experiences the 3→4 transition as a clean demonstration: “here is where the census flips, but here is what remains transportable.” The boundary is not frightening; it is clarifying.

6.4 Why humans experience a “snap” and the CT/HoTT mind does not

The contrast with human experience can now be stated cleanly. Humans typically experience a snap because we bind reality to discrete tokens early. The object-first pipeline commits to a face/edge/vertex decomposition and then treats that decomposition as what the object is. Once that commitment is made, a change in decomposition feels like a change in object identity. This triggers an ontological alarm: “the thing changed category.”

Three specific mechanisms reinforce the snap.

First, early segmentation produces hard boundaries. The mind invests heavily in boundary placement because boundaries are vital for action. Hard boundaries encourage hard incidence. Hard incidence encourages hard counts. Hard counts encourage discontinuity.

Second, labels serve action and communication. We are social animals; we coordinate using discrete terms. “Corner,” “face,” “edge,” “triangle,” “square” are not merely words; they are cognitive handles. When a handle fails, it feels like reality became slippery.

Third, our default feeling of continuity is object-token persistence: the same parts moving smoothly. When parts change, we interpret it as not the same object, unless we consciously override the intuition.

The CT/HoTT mind does not experience a snap because it binds continuity to a different anchor: coherence-preserving transport across presentations. It is not that it lacks discrete tokens; it can produce them when useful. But it does not treat them as metaphysical bedrock. Therefore, when the census of tokens changes, the mind does not infer that being changed. It infers that a chart changed.

In short: humans feel the snap because we locate identity in discrete decomposition; the CT/HoTT mind does not because it locates identity in the existence of witnesses of equivalence—paths of transport and higher compatibilities. The phase change is real at the level of charts, but it is not metaphysically privileged. The mind’s “ordinary” is deeper than the census.

This is the phenomenological heart of the paper. If one accepts that “ordinary” depends on what a mind treats as primary, then the cube–octahedron valence mismatch is not merely a curiosity of polyhedra. It is a window into alternative metaphysics of perception: a world where discrete structure is not denied, but is experienced as the surface of a continuous coherence reality.

7. Ordinary World Competence: Seeing Human Faces

If the cube–octahedron example is a toy problem, faces are the real exam. Human faces are the most socially loaded objects we see, and face perception is among the most refined capacities of the human visual system. If the CT/HoTT-wired mind cannot handle faces, then it is not a plausible mind for our world. Conversely, if the same representational commitments that make polyhedral phase changes feel ordinary also make face perception more robust, then the metaphysics gains credibility: it is not merely an exotic re-description of geometry, but a coherent alternative organizing principle for perception.

This section argues that the CT/HoTT commitments translate naturally into face competence. In fact, faces are a domain where humans already approximate some of these commitments, but usually without explicit access to them as phenomenology. The speculative step is to imagine a mind for which these commitments are not hidden infrastructure but primary "what it is like" to see.

7.1 A face as a type: viewpoint, lighting, expression as internal variation

Humans often talk as if a face is a static shape: a set of features arranged in a particular way. But everyday experience contradicts this. A face changes constantly. It rotates, it moves through light, it occludes itself, it wrinkles and smooths, it smiles, frowns, squints, speaks, blushes, sweats. If a face were represented as a single template, recognition would be brittle. Yet recognition is robust. This suggests that "face" is already, for humans, not a single image but a family of images connected by lawful transformations.

A CT/HoTT-wired cortex would treat this family as the fundamental thing. "Face" would be a type, meaning: a structured space of allowable presentations together with internal coordinates that parameterize lawful variation. Viewpoint, lighting, and expression would not be "nuisance factors" to be removed. They would be dimensions internal to the type.

This reorients the perceptual goal. Instead of "find the canonical face behind the distortions," the goal becomes "locate the current presentation within the face-type and track its coordinates." Put differently: the system does not fight variation; it navigates it.

This has several consequences for what is experienced as ordinary.

First, unusual conditions do not automatically threaten identity. Under object-first assumptions, a strong lighting change can feel like "a different face." Under type-first assumptions, lighting changes are simply movement along a lighting dimension; they are expected internal variation.

Second, partial information remains meaningful. If a face is a type with multiple charts, then even when a chart becomes ill-conditioned (for example, frontal-landmark charts fail under extreme profile), another chart can take over (profile contour chart, ear/cheekbone chart, motion chart). Identity does not vanish because one representation fails; it persists because the type supports multiple presentations.

Third, the mind's sense of sameness becomes graded without becoming vague. The mind can experience: "this is the same person, but with a longer transport," meaning the current presentation is farther away in the type-space from the last stabilized presentation. This is not uncertainty in the everyday sense; it is a felt metric of deformation effort.

7.2 Recognition as transport: identity witnessed by continuous deformation

In object-first cognition, recognition is often experienced as a snap: either you know them or you do not. The snap is useful. It supports quick social action. But it hides the internal mechanism that makes recognition robust: the system is implicitly tracking continuity across time and conditions.

A CT/HoTT-wired cortex makes this continuity explicit at the phenomenological level. Recognition is not "matching." It is transport. The mind experiences identity as witnessed by a continuous deformation that carries one presentation into another while respecting constraints.

This is the perceptual analog of the HoTT idea that equality is not a switch but a path. "This face now" and "that face then" are the same person if there exists a path in presentation-space that links them without violating the face's internal laws.

Several aspects of real faces make the transport framing especially natural.

First, faces are constrained by anatomy. Skin and muscle deformation is not arbitrary. Even without knowing anatomy explicitly, a perceiver can register constraint-consistency: the way a smile pulls cheeks and changes eye shape; the way a jaw opens and stretches the mouth region; the way blinking interacts with gaze shifts. These constraints define the allowable paths. Transport succeeds when the observed changes can be explained as movement along allowable paths.

Second, faces are temporally continuous. In normal life, a face does not teleport between configurations; it evolves. This means identity can be maintained by incremental transport, step by step, rather than by repeated global re-recognition. A CT/HoTT mind would treat this incremental transport as the primary continuity signal.

Third, recognition becomes interpretable. When transport is easy, recognition feels immediate. When transport is hard (poor lighting, occlusion, heavy disguise), recognition feels strained. The mind can locate where the strain comes from: which constraint is difficult to satisfy, which chart is failing, which path is becoming too long. In a human, this is often hidden and only emerges as a vague "not sure." In the CT/HoTT mind, it is experienced as a structured difficulty rather than a binary failure.

This also implies an unusual strength: resistance to superficial changes. Many disguises aim to perturb a few features enough to break template matching. But transport-based identity is harder to fool if the disguise does not preserve the deeper deformation laws and coherence across time.

7.3 Features as subobjects-by-role: eyes and mouth as relational invariants

Object-first face perception tends to treat features as parts: eyes are things, mouth is a thing, nose is a thing. This is often effective, but it can be brittle under occlusion and unusual conditions. A relation-first CT/HoTT cortex would instead treat features as subobjects-by-role: a feature is defined by the relational function it plays within the face-type, not merely by its local appearance.

For example, “an eye” is not just an oval dark region. It is:

- one of a symmetric pair around a midline,
- a region that generates characteristic specular highlights (depending on light),
- a structure whose motion is constrained with eyelids,
- a site whose orientation covaries with head pose and gaze targets,
- a node in a network of relations to brow, cheek, and nose bridge.

Similarly, “the mouth” is not just a curve. It is:

- the primary articulator of speech-related deformation,
- the locus of many expression deformations,
- a structure whose corners are coupled to cheek motion,
- a region whose opening correlates with jaw and chin motion.

When features are defined by role, they remain identifiable across very different visual manifestations. If one cue disappears, others remain: motion, symmetry, relational coupling, temporal dynamics. The feature is not a pixel pattern; it is a stable relational attractor.

This makes the CT/HoTT mind particularly suited to treating face perception as coherence across multiple routes. One route might locate eyes by texture; another by motion; another by symmetry; another by geometry. Agreement across these routes is what makes the “eye” subobject stable. If they disagree, the mind does not simply mislabel; it experiences a coherence problem: the candidate subobjects cannot be reconciled across routes.

This is also where the metaphysical humility of chart changes becomes psychologically important. Under occlusion, the “feature chart” may fail. The mind shifts charts: from texture-based eyes to motion-based eyes, from frontal mouth shape to speech-dynamics mouth role. Identity remains stable because the subobjects are role-defined within a type, not defined by a single appearance template.

7.4 Expression as higher morphism: microexpressions as small homotopies

Expression is the continuous deformation of the face-type over time. In an object-first frame, expressions are often treated as labels: happy, sad, angry, surprised. Those labels are coarse. They compress a rich dynamical phenomenon into discrete categories. Humans can do better than the labels, but we often report labels because language demands them.

In a CT/HoTT-wired cortex, expression would be experienced as a higher-order structure: a morphism between presentations, and microexpressions as small homotopies—tiny deformations that relate one deformation-path to another.

The key idea is that an expression is not just a new face state; it is a transformation of the face state, and that transformation itself can be compared, composed, and related to other transformations. For example:

- A smile can begin abruptly or gradually.
- It can be symmetric or asymmetric.
- It can be genuine or socially performed, with different micro-dynamics.
- It can be interrupted, masked, or blended with another expression.

These are not merely different endpoints. They are different paths. A higher-morphism-aware perceiver would "see" these differences as differences between transformations, not just differences between static shapes.

Microexpressions, on this view, are informative because they are small, fast deformations that reveal underlying constraint structure. They are like tiny perturbations that test the system. If the micro-dynamics are coherent with the person's anatomical and behavioral signature, transport remains easy. If they are not—if the movement is oddly timed, if couplings are inconsistent—transport becomes strained, and coherence begins to fail.

This gives a principled way to talk about the uncanny. The uncanny is not merely "looks fake." It is "the higher morphisms do not cohere." The face may match a static template, but the deformation paths are not compatible with the expected family of paths for that identity. A CT/HoTT mind would therefore be especially sensitive to temporal and relational inconsistencies, because that is where higher-structure lives.

In summary, face perception becomes the everyday analog of the polyhedral story. The "object" (a face) is not a single picture but a type with many charts. Recognition is not a switch but transport along allowable paths. Features are not mere parts but subobjects defined by relational roles. Expressions are not just states but transformations between states, with microexpressions serving as small homotopies that reveal coherence or its failure. If this

architecture sounds nonhuman, it is because it relocates what humans treat as implicit infrastructure into explicit phenomenology: it makes coherence and transport feel as immediate as edges and parts feel to us.

8. Error, Deception, and the Uncanny

Any proposed metaphysics of perception must confront its failure modes. A mind that experiences continuity as transport and "ordinary" as coherence will not be immune to error; it will simply err in characteristic ways. Those characteristic errors matter because they reveal what the mind takes as primitive. In an object-first system, errors often arise from mis-segmentation, mislabeling, or confusing two similar templates. In an equivalence-first CT/HoTT system, errors arise from misplaced coherence: constructing a witness of sameness when none is warranted, or failing to construct one when it is.

This section uses deception and uncanny phenomena as a stress test in the opposite direction from the cube–octahedron example. Earlier we asked: how could a mind avoid overreacting to discrete chart changes? Here we ask: how does such a mind avoid underreacting—how does it prevent "everything is equivalent" from dissolving reality? The answer is that coherence becomes both its power and its danger. Coherence can be faked; witnesses can be forged; transport can be overextended.

8.1 Coherence breaks as immediate perceptual alarms

For an object-first perceiver, the most salient alarms are often boundary-level: a shape is malformed, an edge is wrong, a face looks off, a silhouette is inconsistent. These alarms are powerful, but they can be bypassed by skilled counterfeits that imitate local cues.

For a CT/HoTT-wired perceiver, the primary alarm is different: not "this part looks wrong," but "the parsing routes do not agree." The mind continually attempts to reconcile multiple explanations of the same sensory stream. If the reconciliations fail—if there is no stable way to transport one chart into another while preserving constraints—the system experiences a coherence break.

Phenomenologically, this alarm would feel less like "something is ugly" and more like "something won't compose." The mind tries to integrate cues and experiences resistance: a sense of torsion, mismatch, or shear in representation space. A human analog exists in skilled domains: a musician hears an intonation that clashes not as a wrong note in isolation but as a harmonic incompatibility; an engineer senses that a mechanism "binds" even before identifying the exact part. The CT/HoTT mind generalizes this: compositional incompatibility is felt as perceptual friction.

This changes what counts as “obvious.” A static image that appears plausible might pass an object-first glance, yet fail for a coherence-first perceiver because it does not cohere across time, across viewpoints, or across multiple inferential routes. Conversely, a visually noisy or unusual presentation might still feel real if the coherence witnesses remain constructible.

The important point is that coherence breaks are not post-hoc judgments; they are primary alarms. In the same way that a human sees a sudden motion in peripheral vision and immediately attends, the CT/HoTT mind would feel a breakdown of composability and immediately attend.

8.2 Deepfakes and “failure of lift”: when motion doesn’t match constraints

The phrase “deepfake” names a broad class of synthetic or manipulated media, but our interest is narrower: what would a CT/HoTT-wired mind experience when confronted with a face video that is locally plausible yet globally inconsistent? This is where the earlier notion of “lift succeeds” becomes diagnostic.

Recall the perceptual meaning of lift: the ability to transport a local structure forward through changing conditions while keeping it coherent with the system’s constraints. In face perception, the constraints include anatomy, kinematics, lighting geometry, temporal dynamics of expression, and cross-cue coupling (eyes with lids, cheeks with mouth corners, head motion with background parallax, etc.).

A deepfake can imitate many local statistics. It can produce a face that matches the appearance of a target identity in individual frames. But it is harder to ensure that all constraint couplings remain consistent across time and across all degrees of freedom. The CT/HoTT mind’s alarm would therefore be described not as “I see artifacts,” but as “transport fails”:

- The system attempts to lift identity across frames and finds that the path is not smooth.
- It attempts to reconcile expression dynamics with muscle constraints and finds mismatch.
- It attempts to reconcile lighting changes with 3D geometry and finds a broken link.
- It attempts to reconcile eye motion with head pose and finds noncompositional behavior.

This failure would have a characteristic phenomenology: the face remains recognizable in snapshots, but the identity feels unstable as a continuous entity. It is the difference between recognizing a melody from isolated notes and hearing the melody fall apart when played in time. The “uncanny” would not be primarily a visual strangeness; it would be a temporal-coherence strangeness.

Importantly, a CT/HoTT mind would not be universally skeptical. If the synthetic system successfully preserves the relevant constraints—if it truly constructs a coherent family of deformations—then transport might succeed, and the mind might accept it as real. The criterion is not "natural vs artificial." It is "coherent vs incoherent." In that sense, this mind is metaphysically strict but technologically agnostic: it does not worship biological origin; it worships compositional consistency.

8.3 Disguises as chart tricks: what changes and what cannot

Disguises are a more ancient problem than deepfakes, and they highlight a different vulnerability. A disguise is often a deliberate attempt to force the observer into the wrong chart: to emphasize a presentation that hides role-defining invariants. Glasses, makeup, facial hair, posture changes, clothing, and context manipulation all function as chart shifts.

For an object-first perceiver, many disguises work by perturbing salient parts: change the eyebrows, add a beard, alter the silhouette. Because identity is often tied to a small set of template features, disrupting those features can break recognition.

For a CT/HoTT-wired perceiver, the question becomes: what can a disguise change within the face-type, and what cannot it change without breaking coherence?

A chart trick can alter surface-level cues:

- local texture,
- color distribution,
- apparent contour in certain views,
- even some perceived geometry via padding or prosthetics.

But there are deeper invariants that are harder to fake without a full coherent generative model:

- the coupled dynamics of expression,
- speech articulation patterns over time,
- micro-timing of blinks and saccades,
- the way head motion interacts with shoulders and torso,
- consistent 3D lighting response across motion,
- idiosyncratic asymmetries that persist across many expressions.

In the type-and-transport framework, these deeper invariants are not “features” in the static sense. They are constraints on allowable paths and higher morphisms. A disguise can change the starting point in presentation space, but it cannot easily change the entire path-structure of the identity without becoming, effectively, a full re-embodiment.

Thus the CT/HoTT mind would treat disguises as attempts to force a chart where transport becomes long and awkward. It might initially accept the chart (especially in a brief glance), but as time evolves, the system tries to lift identity across richer dynamics. If the disguise does not preserve the deeper constraint network, the lift strain increases, and coherence begins to crack. The observer experiences a growing sense of “this does not compose,” rather than a simple “I don’t recognize them.”

However, there is also a vulnerability here. Because the CT/HoTT mind is willing to change charts and maintain identity across re-description, it may be more tolerant of superficial changes than a template-matcher. That tolerance is an advantage in benign settings, but it can be exploited: a skilled deceiver might aim not to break recognition completely but to keep the observer in a long, strained transport path where confidence remains low, making social action hesitant.

8.4 Hallucination risk in equivalence-first seeing: false witnesses and overtransport

The most important failure mode of an equivalence-first perception is the one that deserves the name “hallucination” in the strict sense: constructing coherence where none exists. In HoTT language, this would be like accepting a spurious identification path; in categorical language, it would be like forcing a commutative story by smuggling in unjustified transformations. We can call this false witness.

False witness can occur in several ways.

First, overtransport: the mind insists on carrying an identity across too large a gap in presentation space. This is the perceptual analog of seeing a face in clouds or a figure in noise. The system’s hunger for coherence causes it to bridge discontinuities by inventing paths. In humans, this happens as pareidolia. In a CT/HoTT mind, it would happen as an overly generous willingness to declare that a strained transport is still valid.

Second, chart overreach: the mind applies a chart outside its domain of validity. Every chart works well in certain regimes and fails in others. If a mind treats chart transitions as ordinary, it may become complacent about chart discipline and mistakenly trust a chart that is ill-conditioned. The result is a coherent story that is coherent only because the chart is lying—because it is being used where it should not be used.

Third, coherence laundering: the mind uses agreement across multiple routes as its primary reality signal. But routes can be correlated. If the system's parsing routes share the same biases, they can agree for the wrong reasons. This is a deep epistemic risk: "multiple confirmations" may not be independent confirmations. A CT/HoTT-wired cortex would need safeguards against this, analogous to insisting that different routes truly probe different constraints rather than echo the same assumption.

Fourth, narrative completion: a coherence-first system can be tempted to prefer "a complete coherent explanation" over "an incomplete but honest partial." This is the metaphysical version of what, in more mundane terms, is called overfitting. The system would rather generate a witness than admit that no witness is warranted yet. The result is a smooth story that feels ordinary but is ungrounded.

These risks reveal something crucial: equivalence-first seeing is not automatically more truthful. It is more structured, but structure can be misapplied. Therefore, a CT/HoTT mind that is robust must include something like a discipline of disconfirmation: conditions under which it refuses to transport, conditions under which it marks a would-be witness as provisional, and conditions under which it privileges "no identification" over "strained identification."

This is also where our earlier theme returns in a new key. The paper began with the idea that humans overreact to discrete chart changes and experience snaps. Here we see the opposite danger: a mind so fluent in chart changes that it underreacts to genuine discontinuity and invents continuity. The right stance is not "always smooth." It is "smooth when warranted by coherence witnesses, discrete when the witnesses cannot be constructed."

In summary, deception and the uncanny are not special cases. They are a mirror held up to a perceptual metaphysics. Object-first seeing is brittle to superficial part-perturbations but often conservative about identity across radical changes. Equivalence-first seeing is robust to superficial variation and comfortable with re-description, but it is vulnerable to false witnesses, route correlation, and overtransport. Any serious account of a CT/HoTT-wired cortex must include not only how it achieves ordinary continuity, but also how it refuses counterfeit continuity when the world does not support it.

9. Cognitive Tradeoffs and Limits

A mind that experiences "ordinary reality" as coherence-preserving transport has obvious strengths: it can absorb chart changes without panic, it can track identity through variation, and it can detect certain kinds of inconsistency that evade part-based heuristics. But no perceptual metaphysics is free. If we imagine a CT/HoTT-wired cortex as a genuine biological or engineered mind, it must pay for its gifts. The payment is not only computational; it is also existential and

social. A world that is parsed primarily as types, transports, and witnesses will not feel the same as a world parsed as stable objects with parts.

This section is therefore a counterweight. It lists the costs, the social implications, the necessity of discretization for action and communication, and the characteristic pathologies that would arise when coherence-first perception is pushed beyond its proper domain.

9.1 Costs of coherence-first perception: time, compute, and attentional load

Coherence-first perception is expensive because it refuses to settle early. It maintains multiple charts, multiple parsing routes, and multiple candidate transports until coherence resolves them. In object-first perception, early commitment is a feature: decide quickly, act quickly, revise only when forced. Coherence-first perception reverses that priority: keep multiple candidates alive, evaluate their compatibility, and choose the most coherent representation. That is compute.

Several concrete costs follow.

First, time cost. If you test coherence across routes, you must run the routes. If you attempt transport, you must attempt it repeatedly over time. The system's experience of continuity is not free; it is paid for by continuous verification. This implies that the system may have a slightly slower "snap-to-label" response, especially in ambiguous settings. Where a human might quickly declare "that's him" or "that's not him," a coherence-first mind might remain in a state of "transport in progress" longer.

Second, compute cost. Tracking multiple charts and routes resembles maintaining a small ensemble of models rather than a single model. Even if the underlying machinery is efficient, the logical requirement remains: coherence is defined by agreement of multiple compositional paths. That agreement cannot be computed without maintaining alternative paths.

Third, attentional load. Attention is not just selecting a region of the visual field; it is selecting which hypotheses to keep alive and which to prune. A coherence-first system has a different attention problem: it must allocate attention to the places where coherence is most informative. Those places are often boundary events and transition zones—regions where charts become ill-conditioned and re-factorization is imminent. Such zones may attract attention repeatedly. The mind may develop a "critical-event bias," habitually drawn to regime shifts, anomalies, and subtle inconsistencies.

Fourth, energy and fatigue. If coherence is continuously evaluated, the system may suffer a kind of perceptual fatigue in highly noisy or adversarial environments, where transports often fail and charts continuously compete without resolution. Humans experience something similar when trying to follow many speakers in a loud room: the effort is draining. A coherence-first

mind might be more drained by modern environments saturated with synthetic, edited, and attention-engineered stimuli.

Finally, there is a meta-cost: the need for stopping rules. A system that always seeks a witness of equivalence can get stuck searching. Action requires termination. So the architecture must include explicit "good enough" thresholds: when to commit, when to suspend, when to ignore. Those thresholds are not merely engineering details; they shape phenomenology. The mind's ordinary world depends on its ability to stop seeking perfect coherence.

9.2 Social metaphysics: individuality as a section through a continuous space

Perhaps the deepest tradeoff is not computational but social. A coherence-first mind, by design, lives in continuous spaces of variation: types with families of presentations. For faces, this means that "personhood in perception" is naturally represented as a region in a high-dimensional space, not as a single point. Identity is a bundle of allowable deformations, not a static template.

This changes the feel of individuality.

In object-first social perception, individuality is crisp: that person is that person. Even when we are wrong, we experience the crispness. We experience persons as discrete units. A coherence-first mind can still produce discrete identities for practical purposes, but it also sees the underlying continuity: family resemblance, genetic similarity, shared mannerisms, convergent appearance under certain lighting, drift over time with aging. It may experience identities as neighboring basins in a continuous landscape.

The paper's earlier language of "a section through a continuous space" becomes socially charged here. A section is a choice of a representative from each equivalence class, a way of picking a discrete handle for something that is, underneath, continuous. Social life requires such sections: names, categories, roles. But a mind that experiences the continuity more directly may experience the section as conventional rather than metaphysically absolute.

This can have both humane and alien effects.

On the humane side, such a mind might be less prone to harsh categorical judgments based on superficial appearance because it sees appearance as variation within types and neighborhoods. It might perceive more continuity between individuals, more shared structure, more "ways this person could be, given the space they inhabit."

On the alien side, it might find certain social practices strange: rigid identity policing, fixation on superficial tokens, obsession with hair/skin/angle-based aesthetics. It might also struggle emotionally if it cannot, by default, experience persons as crisp discrete absolutes. Human attachment often feeds on discrete recognition: "you, not another." A mind that perceives

strong continuity among near-neighbors might need additional mechanisms to preserve the felt sanctity of individual persons, lest persons become “points in a manifold” rather than morally irreducible beings.

This is not a claim that such a mind would be unethical. It is a warning that its perceptual metaphysics would pressure social metaphysics. It would need explicit commitments—perhaps moral or covenantal—ensuring that continuous perceptual spaces do not dissolve the inviolability of individuals.

9.3 When discretization is necessary: action, language, and shared reference

Even if coherence-first perception is metaphysically elegant, the world demands discrete acts. You must step here or there. You must grasp this object now. You must say a name. Discretization is not merely a cognitive convenience; it is required by embodiment and communication.

Three necessities enforce discretization.

First, action requires crisp boundaries. A hand must close around a region; a foot must land. Biological control loops cannot operate on indefinite charts forever. They must commit to a segmentation that supports motor planning. Therefore even a CT/HoTT cortex must include a mechanism that produces hard boundaries on demand—perhaps as a downstream projection from the richer coherence state.

Second, language requires shared reference. Communication depends on common tokens. If one mind experiences identity as a family of transports, but the community exchanges names and categories, the mind must choose a section: it must map the rich internal structure to a label that others will understand. This means that, socially, the mind will appear more ordinary than it is. Its internal world may be continuous and witness-based, but its outputs must be discretized into culturally stable symbols.

Third, moral responsibility requires commitments to discrete claims. In many contexts—testimony, accusation, promise—one must say: yes or no; this person or not; this event happened or did not; this is the same contract or not. A coherence-first mind might be tempted to remain in graded uncertainty, but the world sometimes demands a decision. So the architecture must incorporate decision policies that convert graded coherence into discrete commitments, while also preserving the ability to retract and revise when new coherence evidence arrives.

This is where “chart change without rupture” must be disciplined. If every change is experienced as re-description, the mind could become too tolerant of revision, undermining stable commitments. Therefore discretization is not merely a practical layer; it is a moral and

social necessity. A responsible coherence-first perceiver must know when to stop treating everything as a re-description and instead treat some boundaries as binding for action and accountability.

9.4 Pathologies: overidentification, undersegmentation, and missed saliencies

Finally, every perceptual metaphysics has signature pathologies. For object-first vision, common failures include pareidolia at the part level (seeing faces in clouds), brittle misrecognition under occlusion, and overconfidence in hard boundaries when the world is ambiguous. For coherence-first vision, the pathologies are different.

First, overidentification. This is the tendency to construct a transport witness too readily: to declare two things "the same" because a strained path exists, even when the path should not be accepted. Overidentification can manifest as false recognition ("that's him") or as unwarranted continuity ("it's still the same situation") in dynamic environments. In human terms, it resembles seeing patterns where there are none, but in a structured way: not random pareidolia, but coherence-driven pareidolia.

Second, undersegmentation. Because the system prefers continuity and is comfortable with chart changes, it may delay or resist splitting one entity into two when the evidence demands it. For example, in a crowded scene with similar faces, a coherence-first system might be reluctant to break continuity and might mistakenly transport identity across two different individuals. This is a failure of identity separation, not a failure of feature detection.

Third, missed saliencies. Human object-first vision is tuned to sharp boundaries, novelty, and sudden motion. A coherence-first system might instead prioritize global consistency. In doing so, it might miss small but crucial local details that do not strongly affect coherence. A pickpocket's subtle hand motion might be locally salient but globally coherent. A tiny object (a key) might be crucial for action but not strongly constrained by the coherence network. Thus, coherence-first perception might require auxiliary saliency mechanisms—modules that flag locally important details even when they do not threaten global consistency.

Fourth, coherence laundering and correlated routes. As discussed earlier, agreement across routes is only reliable if routes are independent. If the system's routes share biases, it can be tricked into stable but wrong coherence. This pathology could be especially dangerous in social contexts: the system might become confidently wrong in a way that feels deeply ordinary because coherence is high. The mind would need explicit diversity in its parsing routes and explicit skepticism about agreement that arises from shared assumptions.

Fifth, paralysis by coherence. In complex, adversarial, or highly ambiguous environments, coherence may never settle. If the mind insists on witnesses before acting, it can become indecisive. Therefore healthy function requires stopping rules, trust heuristics, and hierarchical

fallbacks: when coherence is too expensive, act on coarse discretizations and accept higher error risk.

These tradeoffs and pathologies do not refute the CT/HoTT perception hypothesis. They complete it. A plausible mind is not defined only by what it can do; it is defined by how it fails. The distinctive promise of coherence-first perception is that it can make combinatorial "phase changes" ordinary by relocating continuity from discrete censuses to transportable structure. The distinctive danger is that it can blur the moral and practical necessity of discreteness by overextending equivalence. The balance between these forces—between transport and commitment, between chart flexibility and decision discipline—is where such a mind would live.

10. Testable Predictions and Thought Experiments

A metaphysics of perception earns its keep when it yields discriminators: outcomes that would look different under object-first versus coherence-first (CT/HoTT-like) seeing. The claims in this paper are not merely poetic. They imply measurable signatures in behavior, error patterns, confidence dynamics, and sensitivity to particular classes of perturbations. This section proposes four families of tests. Each can be run as a human psychophysics experiment, as an evaluation suite for vision models, or as a thought experiment that clarifies what would count as "the same mind" across architectures.

10.1 Behavioral signatures: invariance profiles under occlusion and lighting

The central prediction is not "better recognition overall," but a distinctive *shape* of invariance: what breaks recognition, what merely lengthens transport, and what triggers a coherence alarm.

A practical way to test this is to measure recognition accuracy and reaction time as a function of controlled perturbations:

- Occlusion sweeps: progressively occlude the face (random masks, structured masks, lower-half/upper-half masks) while keeping other variables stable.
- Lighting sweeps: vary directionality, contrast, and color temperature systematically.
- Viewpoint sweeps: rotate yaw/pitch/roll in controlled steps.
- Dynamic sweeps: compare static frames versus short clips with natural motion.

Object-first prediction: performance degrades sharply when key parts (eyes/mouth) are occluded or when canonical templates fail; confidence often remains high until it suddenly collapses (a "cliff" profile).

CT/HoTT prediction: performance degrades more gracefully under many superficial perturbations (a "slope" profile), because identity is maintained by chart-switching and transport across cues. However, the CT/HoTT mind should show an additional signature: *increased reliance on time* as a stabilizer. Under harsh occlusion or lighting, it may require longer clips to succeed, because transport needs sequential constraint evidence to construct a stable witness.

A concrete discriminator is the "time-to-stability" curve: how long it takes for a perceiver's identity judgment to settle as perturbations increase. Object-first systems tend to settle quickly or fail quickly. Coherence-first systems tend to settle more slowly but remain correct longer—until they hit genuine coherence failure modes.

10.2 Ambiguity handling: multiple identifications retained rather than collapsed

A coherence-first mind is predicted to treat ambiguity as a structured intermediate state, not an error to be eliminated immediately. The behavioral signature is *non-collapse*: it keeps more than one identification live, with graded support, and resolves later when coherence witnesses accumulate.

This can be tested by changing the response format. Instead of forcing a single identity choice, allow:

- Multiple candidates (select up to k identities).
- A graded distribution (assign weights summing to 1).
- A "transport strain" rating (how hard the identification feels).
- A "chart shift" report (did your basis for recognition change: texture → motion → contour, etc.).

Then present sequences designed to remain ambiguous for a while and then disambiguate (e.g., occluded face becomes visible; profile rotates to frontal; lighting changes reveal 3D cues).

Object-first prediction: participants often commit early to one identity and then reverse abruptly when evidence contradicts it (high flip rate, high overconfidence, sharp reversals).

CT/HoTT prediction: participants (or models) will show delayed commitment, broader candidate sets early, fewer hard reversals, and smoother redistribution of weights—because identity is maintained as "transport in progress" rather than as an immediate label.

A strong discriminator is the "reversal geometry": do judgments change by abrupt flips, or by continuous reweighting? If the CT/HoTT metaphysics is even approximately correct, coherence-

first systems should exhibit continuous reweighting as the default mode, with flips occurring mainly when coherence truly breaks.

10.3 Detection of synthetic faces: constraint-consistency sensitivity

The paper predicts that the uncanny is not primarily a static artifact but a *failure of lift*—a breakdown in the ability to transport identity and local structure across time while satisfying coupled constraints.

This suggests an evaluation suite that is not "real vs fake" in the broad sense, but "constraint-consistent vs constraint-inconsistent" while holding superficial realism constant.

Construct stimuli in controlled families:

- Kinematic inconsistencies: eyelid motion decoupled from eye rotation; mouth corners move without cheek response; jaw motion inconsistent with chin/neck coupling.
- Lighting/geometry inconsistencies: specular highlights that do not move consistently with head rotation; shading that contradicts implied 3D structure.
- Temporal inconsistencies: microexpression timing that violates typical smoothness; blink cadence inconsistent with gaze shifts; audio–lip timing mismatches.

Object-first prediction: detection correlates strongly with local artifacts and frame-level anomalies; highly polished deepfakes pass more often.

CT/HoTT prediction: detection correlates with *cross-cue coupling failures* and *temporal inconsistency*, even when individual frames look plausible. The system may accept degraded-looking but coherent footage more readily than polished but incoherent footage.

A concrete discriminator is "coherence-first AUC": train or select perturbations that preserve static realism but violate coupling laws, and measure sensitivity. Coherence-first perceivers should outperform object-first ones specifically on these "high realism / low coherence" cases.

10.4 The "polyhedron movie" task: graded incidence perception in humans and models

To test the original cube–octahedron thesis directly, we need a task that does not reduce the phenomenon to a single forced-choice ("cube or octahedron?"). We need to measure whether a system can perceive *graded incidence* and treat the 3→4 change as a chart transition rather than a metaphysical snap.

Design a "polyhedron movie" stimulus: a continuous animation over a parameter t in [0,1] that moves through a family of shapes where a discrete polyhedral census changes. The exact geometric family can be implemented in many ways (implicit surfaces with sharpening,

controlled truncations, or duality-guided transitions), but the test is invariant: we are measuring how observers represent the transition.

Ask participants (or models) to provide time-series judgments while watching:

- Incidence strength at each candidate vertex: “How strongly do you perceive a fourth face participating here?” (a continuous slider).
- Chart confidence: “How well does the ‘3-face vertex’ description fit right now?” and separately “How well does the ‘4-face vertex’ description fit right now?”
- Transport continuity: “Does the identity of this local junction feel continuous?” (yes/no plus strain rating).
- Boundary localization: “Mark the moment you believe the combinatorial description must change.”

Object-first prediction: incidence strength remains near 0 until near the boundary, then jumps; chart confidence switches abruptly; boundary time is sharply localized; participants report a snap-like moment.

CT/HoTT prediction: incidence strength ramps gradually; both chart confidences overlap in a transition zone; boundary time is spread (a region, not a moment); participants report continuity of the junction with a localized region of “chart strain” rather than rupture.

For models, there is a parallel internal diagnostic: does the representation evolve continuously with a detectable “transition zone” (overlap of competing parses), or does it abruptly reclassify? A CT/HoTT-like model should show a measurable overlap regime where two decompositions are both partially supported and a higher-level consistency signal selects the chart without breaking transport.

The deeper point of this task is that it operationalizes the paper’s metaphysical claim. “Seeing the 3→4 switch smoothly” is not a mystical talent. It is a measurable pattern: graded participation, overlap of charts, preserved transport, and boundary crossings experienced as ordinary. If any human subgroup, training regime, or model architecture exhibits these signatures, that would be evidence that the CT/HoTT perceptual metaphysics is not merely science fiction but a real, discoverable mode of seeing.

11. Implications for AI, XR, and Neuroscience

If the CT/HoTT-perception hypothesis were only a metaphysical game, it would remain safely untestable. But it points toward concrete design principles and research questions. It suggests

ways to build perception systems that are resilient to regime shifts and chart changes, and it suggests new behavioral assays for how humans stabilize identity under ambiguity. It also raises ethical concerns because a coherence-first perceiver, whether human-trained or machine-built, could become unusually good at certain kinds of recognition and inconsistency detection—precisely the capacities that can be weaponized.

This section organizes implications across three domains: AI architecture, XR as a training environment, and neuroscience as a measurement discipline. We end with ethical constraints, because "better seeing" is never only a technical win.

11.1 Building perception stacks that are functorial by design

Most deployed vision systems are pipeline-like in a way that is *not* functorial. They map input images to labels or to bounding boxes, often with intermediate representations that are useful but not explicitly constrained to preserve compositional structure. "Functorial by design" means something more principled: the system should preserve the way relations compose when translating from sensory streams to inferred world structure.

What would that look like in practice?

First, treat transformations as first-class objects. Instead of only predicting "what is in the frame," the system predicts "how the scene is changing" as a structured morphism: pose change, illumination change, deformation field, camera motion, object articulation. Then it composes these morphisms over time. A perception stack becomes a composition engine: the identity of a face is stabilized not by repeated classification, but by the existence of a coherent transport chain across frames.

Second, enforce multi-route agreement explicitly. Rather than a single encoder feeding a single decoder, the system maintains multiple parsing routes that use different cues (texture, geometry, dynamics, context). The goal is not redundancy for its own sake, but *coherence scoring*: the system learns to value representations that reconcile across routes, and to treat irreconcilable disagreement as an alarm.

Third, include chart-change mechanisms. Real data is full of regimes: frontal vs profile faces, day vs night scenes, occluded vs unoccluded objects. A functorial stack should not pretend one representation works everywhere. It should maintain chart families and transition logic: when a chart becomes ill-conditioned, the system should switch in a controlled way while preserving identity via a transport witness.

Fourth, make "lift succeeds" an internal objective. Instead of optimizing only for end labels, optimize for the ability to transport structure across time and perturbations. This can be implemented as a consistency objective: if the system predicts a structure at time t and a

transformation to time t+1, then applying the transformation to the structure should match the inferred structure at t+1. This is not merely a self-supervised trick; it is the operational definition of transport.

Fifth, build explicit stop rules and skepticism gates. As discussed earlier, coherence-first systems can overtransport. A functorial stack must know when to refuse a witness. That suggests training for calibrated "no identification" outcomes and penalizing confident wrong transports more than uncertain suspensions.

The broader implication is that "category theory" here is not a library of formal symbols but a discipline: treat perception as composition, preserve structure across mappings, and value coherent commutativity across routes. The design principle is: the model should behave as if it respects diagrams, even if it never manipulates explicit diagrams.

11.2 Representation choices: meshes versus implicit fields versus path-based models

The cube–octahedron problem highlights that discrete decompositions are brittle under phase changes. In AI terms, this is a representation problem: what internal format can remain continuous when a discrete census (like face count at a vertex) must change?

Three broad representation families matter.

First, mesh-based representations. Meshes (explicit vertices/edges/faces) are powerful for rendering and for many geometric tasks, but they encode the discrete decomposition explicitly. When the decomposition must change (edge collapse, face split), the representation undergoes a topological event. That event can be handled, but it is naturally "snappy" unless the system includes explicit machinery for remeshing and maintaining identity across remesh operations. Mesh-based systems align naturally with object-first cognition: parts are explicit.

Second, implicit-field representations. An implicit surface or signed distance field can deform continuously as a field, and discrete faces and edges can be extracted later by thresholding or fitting planar patches. This aligns with the paper's idea that valence is a derived census. Implicit fields naturally support "soft incidence": the junction structure emerges from gradients and curvature regimes rather than being hard-coded. However, implicit fields alone do not guarantee coherence-first behavior; they merely provide a continuous substrate on which coherent charts could be defined.

Third, path-based or transport-centric representations. Here the primary representational object is not a static shape, but a *family of shapes connected by transformations*. In practice, this means the system stores not only geometry but the admissible deformations and the witnesses that connect different presentations. The representation is inherently temporal and relational: identity is a transport invariant. This is closest to the HoTT flavor: identity is the

existence of a path in representation space. It is also the hardest to implement and evaluate, because it shifts the target from “reconstruct the object” to “maintain coherent equivalence across views.”

The key insight is that no single representation is sufficient. A robust coherence-first stack likely needs a hybrid: implicit fields for continuity, mesh-like summaries for action and communication, and path-based witnesses for identity across re-factorizations.

This hybrid framing also clarifies why the cube–octahedron transition is a good diagnostic. It forces the system to handle a regime where mesh topology changes, and it asks whether the system can maintain a stable identity witness across that change. If it can, it is closer to relation-first seeing.

11.3 XR as training for chart-change perception: can humans learn this qualia?

A striking implication of the paper is that “ordinary” depends on primitives, and primitives can sometimes be trained. XR (VR/AR) is therefore not just entertainment; it is a laboratory for perceptual metaphysics.

The question is whether humans can learn to experience chart changes as non-rupturing—whether the “snap” can be softened by training that encourages maintaining identity across re-descriptions.

XR provides tools that ordinary life does not:

- You can control the exact transformation family of a stimulus.
- You can slow down or speed up transitions to expose the overlap zone.
- You can provide interactive control (users can “scrub” the parameter t).
- You can overlay alternative charts simultaneously (primal and dual views).
- You can provide feedback on coherence, not just correctness.

A plausible training protocol would not try to force people to “see faces as types” by explanation. It would shape experience by rewarding transport. For example:

- Present “polyhedron movies” where the participant’s task is to track a specific junction continuously and report graded participation, not to label shapes.
- Provide dual overlays: show primal faces and dual vertices simultaneously, letting the user learn co-presence.

- Use haptic or auditory feedback as a coherence signal: when the participant's tracking is transport-consistent, the system produces a smooth tone; when their tracking jumps, the tone becomes noisy.
- Gradually increase complexity: from cube↔octahedron to more complex dual pairs and then to face-like objects.

The prediction is not that humans will become alien. It is that some aspects of the CT/HoTT phenomenology might be learnable: increased tolerance for chart changes, reduced snap-based reporting, and improved ability to retain multiple candidate parses through ambiguity.

If such training works even partially, it would be significant. It would imply that "CT/HoTT-like seeing" is not purely nonhuman; it is a latent mode that can be cultivated, at least as a higher-level perceptual skill. If it does not work—if people revert to snaps no matter what—this would suggest that object-first primitives are deeply constrained by human neural architecture.

Either outcome is informative.

11.4 Ethics: enhanced recognition, privacy, and the temptation of "oracle sight"

Any system—human-trained or machine-built—that excels at coherence-based identity tracking raises ethical risks. The risks are not limited to surveillance. They extend to social power, misinterpretation, and a subtle spiritual temptation: to treat enhanced perception as an oracle.

Several ethical points follow directly from the paper's metaphysics.

First, enhanced recognition threatens privacy. If a coherence-first system can recognize people through disguises, lighting extremes, and partial occlusions by relying on deep transport invariants, then anonymity becomes harder. The ethical response cannot be "do not build it" in all contexts, because medical and safety contexts may benefit. But it does require strict governance: purpose limitation, access control, audit trails, and strong constraints on deployment.

Second, coherence alarms can be abused. A system that flags "constraint inconsistency" can be used to accuse: "your video is fake," "your testimony is incoherent," "your face does not match your identity." Such claims can carry undue authority. Yet coherence is not truth; it is consistency with a model. A coherent lie is possible. An incoherent truth is possible under noise, trauma, disability, or atypical physiology. Therefore coherence tools must be framed as probabilistic and contestable, not as verdict machines.

Third, the temptation of "oracle sight." The phenomenology described in this paper is seductive: a mind that feels coherence breaks as immediate alarms and transport success as deep continuity could easily slide into overconfidence. It might treat its own perceptual coherence as

a proxy for metaphysical certainty. That is a cognitive and moral risk. In human terms, it is the temptation to worship one's own discernment.

Any responsible design—whether training humans in XR or building AI systems—must include humility rules: explicit uncertainty reporting, refusal to overtransport under adversarial conditions, and institutional checks that prevent coherence judgments from becoming unchallengeable authority.

Finally, there is a dignity constraint. Faces are not merely signals; they are persons. A perception technology that treats faces as richly navigable type-spaces may inadvertently encourage objectification: people become trajectories in a manifold. To counter this, we need an explicit moral stance: whatever the representation, the person is not reducible to it. Enhanced perception should increase reverence, not domination.

In summary, the implications are double-edged. AI can be made more robust by building functorial stacks that preserve structure and treat identity as transport. Representation choices matter: continuous substrates plus explicit witnesses can tame phase-change snaps. XR might train humans to tolerate chart changes and perceive graded incidence. But these same capacities threaten privacy and invite oracle-like overreach. If we pursue coherence-first seeing, we must pair it with governance: auditability, contestability, and humility—so that "better seeing" does not become a new way to harm.

12. Conclusion: A Different Ordinary

The cube–octahedron puzzle looks small: a child can learn that three faces meet at a cube vertex and four meet at an octahedron vertex. Yet the perceptual resistance is real. Many of us can watch a smooth morph between polyhedra and still feel that, at some point, something "snaps" in the ontology. The moral of this paper is that the snap is not a fact about geometry alone. It is a fact about what a mind treats as primitive.

We have used the cube–octahedron duality as a stress test for metaphysics of perception. The result is a proposal: if a mind's primitives are compositional, functorial, and path-based—if it treats coherence and transport as primary—then the 3→4 switch becomes a normal transition. Not because discreteness disappears, but because discreteness is demoted from bedrock to readout.

12.1 Restating the thesis: phase changes are chart changes under coherence-first vision

The thesis can be stated in one sentence: what humans experience as a discrete phase change is, for a coherence-first (CT/HoTT-like) perceiver, a chart change inside a continuous structure.

More explicitly:

- The underlying percept is a type: a structured family of allowable presentations.
- "Faces," "edges," "vertices," and "valence" are not metaphysical atoms; they are chart-level summaries extracted from that family when a particular representation is well-conditioned.
- As the system moves through a boundary region (a stratified locus), one chart becomes strained and another becomes stable.
- Identity persists because it is anchored not in a fixed decomposition but in the existence of coherence-preserving transports—witnesses that carry local structure across re-factorizations.

Under this metaphysics, "3 faces vs 4 faces" is a census difference, not a metaphysical rupture. The discrete count changes because the best summary changes, while the underlying structure remains continuously followable. The feeling of continuity is supplied by "lift succeeds": the mind can transport the local junction through the transition without losing coherence. The transition is therefore experienced as ordinary: a moment of representational refactoring, not a moment of creation or annihilation.

This is the sense in which we have proposed "a different ordinary." Ordinary is not what is simple; it is what does not threaten being.

12.2 Cube–octahedron duality as a parable for everyday recognition

If the argument stopped at polyhedra, it would be a curiosity. The deeper claim is that the same machinery that makes polyhedral phase changes ordinary also makes the ordinary world more legible—especially the world of faces.

A face, on this paper's account, is already a type: viewpoint, lighting, expression, age, and occlusion are internal variations, not external annoyances. Recognition is transport: identity is witnessed by continuous deformation through lawful constraints. Features are subobjects-by-role: eyes and mouth are stable not merely as image patches, but as relational invariants across multiple routes. Expression is higher structure: microexpressions are small homotopies that reveal whether the deformation paths cohere.

Seen this way, the cube–octahedron story becomes a parable. The parable is that much of what we treat as "identity" in daily life is already maintained across chart changes. We do not recognize a loved one because one static template matches; we recognize them because the world supplies coherent transports across time and variation. Our human snap is therefore not

proof that reality is snap-like. It is proof that we often compress continuity into discrete handles early, for speed and action.

The CT/HoTT mind is not a magical seer. It is a mind that refuses to bind ontology to the handles. It binds ontology to the witnesses.

12.3 Forward work: formalization boundaries, experiments, and responsible deployment

The forward agenda has three tracks: how far to formalize, how to test, and how to govern.

First, formalization boundaries. It is tempting to turn this paper into a full formal program—categories of perceptual states, functors from sensation to structure, homotopies as equivalences between parsing paths. But formalism can become a substitute for insight. The appropriate boundary is functional: formalize only what yields discriminators and design principles. The goal is not to claim "the brain is doing HoTT," but to specify what computational commitments would reproduce the proposed phenomenology.

Second, experiments. The paper proposed concrete tests: invariance profiles under occlusion and lighting, ambiguity retention rather than early collapse, sensitivity to constraint-consistency failures in synthetic faces, and the "polyhedron movie" task measuring graded incidence perception and overlap zones between charts. These tests can be run on humans (including trained subgroups), on modern vision models, and on hybrid systems. The objective is to detect not just accuracy differences but *signature dynamics*: continuous reweighting versus discrete flips, transport strain versus label failure, and boundary zones experienced as ordinary versus rupturing.

Third, responsible deployment. If coherence-first perception can be engineered, it will sharpen recognition and anomaly detection. That is valuable in medicine, safety, and accessibility. It is also dangerous in surveillance, coercion, and "oracle politics." The ethical requirement is therefore not optional. Any deployment must include purpose limitation, auditability, contestability, calibrated uncertainty, and humility rules that prevent coherence judgments from becoming unchallengeable authority. The more a system feels like it "sees through" appearances, the more it must be constrained to avoid becoming a tool of domination.

The final claim of the paper is modest but consequential: the hardest part of the cube–octahedron transition is not the geometry. It is our choice of perceptual primitives. If we imagine a mind whose primitives are compositional and path-based, the 3→4 switch becomes something like a change in coordinate system—noticeable, even beautiful, but not metaphysically alarming. That mind would live in a different ordinary: one where discrete structure is real, yet not ultimate; where continuity is not the absence of thresholds, but the presence of witnesses; and where the deepest percept is not "what parts are there," but "what coherence can be carried forward."

References

- Awodey, S. (2010). *Category Theory* (2nd ed.). Oxford University Press.
- Baez, J. C., and Stay, M. (2009). Physics, topology, logic and computation: A Rosetta stone. arXiv:0903.0340. (Also appears in *Lecture Notes in Physics* 813, 2011.)
- Bruce, V., and Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77(3), 305-327.
- Coxeter, H. S. M. (1973). *Regular Polytopes* (3rd ed.). Dover Publications.
- Dolhansky, B., Bitton, J., Pflaum, B., Lu, J., Howes, R., Wang, M., and Canton Ferrer, C. (2020). The DeepFake Detection Challenge (DFDC) dataset. arXiv:2006.07397.
- Fong, B., and Spivak, D. I. (2019). *An Invitation to Applied Category Theory: Seven Sketches in Compositionality*. Cambridge University Press.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11, 127-138.
- Giaquinto, M. (2007). *Visual Thinking in Mathematics: An Epistemological Study*. Oxford University Press.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin.
- Goresky, M., and MacPherson, R. (1988). *Stratified Morse Theory*. Springer.
- Grunbaum, B. (2003). *Convex Polytopes* (2nd ed.). Springer.
- Hansen, J., and Ghrist, R. (2019). Toward a spectral theory of cellular sheaves. *Journal of Applied and Computational Topology*, 3, 315-358. (Preprint: arXiv:1808.01513.)
- Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223-233.
- Helmholtz, H. von. (1867). *Handbuch der physiologischen Optik* (foundational "unconscious inference" tradition; various later editions/translations).
- Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11), 4302-4311.
- Knill, D. C., and Richards, W. (eds.). (1996). *Perception as Bayesian Inference*. Cambridge University Press.
- Leinster, T. (2014). *Basic Category Theory*. Cambridge University Press.

- Lee, J. M. (2013). *Introduction to Smooth Manifolds* (2nd ed.). Springer.
- Lurie, J. (2009). *Higher Topos Theory* (Annals of Mathematics Studies 170). Princeton University Press.
- Mac Lane, S. (1998). *Categories for the Working Mathematician* (2nd ed.). Springer.
- MacDorman, K. F., and Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, 7(3), 297-337.
- Mirsky, Y., and Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, 54(1), Article 7.
- Mori, M. (1970). The uncanny valley. *Energy*, 7(4), 33-35. (Often cited via later translations/reprints.)
- Pflaum, M. J. (2001). *Analytic and Geometric Study of Stratified Spaces*. Springer.
- Phillips, S., and Wilson, W. H. (2010). A category theory explanation for the systematicity of human cognition. *PLoS Computational Biology*, 6(7), e1000858.
- Phillips, S. (2022). What is category theory to cognitive science? *Topics in Cognitive Science* (and/or related journal versions; see PubMed record).
- Rao, R. P. N., and Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79-87.
- Riehl, E. (2017). *Category Theory in Context*. Dover Publications.
- Robinson, M. (2016). Sheaves are the canonical data structure for sensor integration. arXiv:1603.01446. (Expanded journal version appears in *Information Fusion*, 2017.)
- Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., and Nießner, M. (2019). FaceForensics++: Learning to detect manipulated facial images. *ICCV 2019*. (Preprint: arXiv:1901.08971.)
- Shin, S.-J. (1994). *The Logical Status of Diagrams*. Cambridge University Press.
- Spivak, D. I. (2014). *Category Theory for the Sciences*. MIT Press.
- Thies, J., Zollhofer, M., Stamminger, M., Theobalt, C., and Nießner, M. (2016). Face2Face: Real-time face capture and reenactment of RGB videos. *CVPR 2016*.

- Thies, J., Zollhofer, M., and Nießner, M. (2019). Deferred neural rendering: Image synthesis using neural textures. *ACM Transactions on Graphics* (SIGGRAPH 2019). (Preprint: arXiv:1904.12356.)
- The Univalent Foundations Program. (2013). *Homotopy Type Theory: Univalent Foundations of Mathematics*. arXiv:1308.0729.
- Voevodsky, V. (2014). *Univalent Foundations* (IAS / IHP lecture slides and notes; various archival versions).
- Ziegler, G. M. (1995). *Lectures on Polytopes*. Springer.

The author has published over 4,700 AI-assisted papers at: https://www.researchgate.net/profile/Douglas-Youvan/research . Google Scholar has indexed these preprints, and if you want a particular reference, the AI mode of a Google search (Gemini) has efficiently catalogued these preprints.